QUORUM SENSING, ENTRY EXCLUSION AND REPLICATION OF THE TI PLASMID OF *AGROBACTERIUM TUMEFACIENS*

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Uelinton Manoel Pinto

May 2011

QUORUM SENSING, ENTRY EXCLUSION AND REPLICATION OF THE TI PLASMID OF *AGROBACTERIUM TUMEFACIENS*

Uelinton Manoel Pinto, Ph. D.

Cornell University 2011

Agrobacterium tumefaciens contains the tumor inducing (Ti) plasmid that stimulates tumor formation in wounded plant tissues. Replication, partitioning, conjugation and entry exclusion of this plasmid are under strict cellular control, and are regulated by quorum sensing. In Gram-negative bacteria, the quorum sensing circuitry is usually mediated by small diffusible signal molecules produced by LuxI homologues and signal receptor proteins homologous to LuxR. In *A. tumefaciens*, TraR is the LuxR-type quorum sensing receptor that is activated by the signal molecule N-3-oxooctanoyl-L-homoserine lactone (OOHL). We studied the TraR dimerization properties, the involvement of quorum sensing on entry exclusion, and replication of the Ti-plasmid.

First we wanted to know whether TraR dimerization is required for protein stability. It is well established that transcriptional activation by TraR depends on the presence of the signal molecule OOHL, which is required for protein folding, protease resistance, and dimerization. My results suggest that dimerization of TraR enhances resistance to cellular proteases, further contributing to protein stability and function.

In a second study it is shown that the Ti plasmid encodes for robust entry exclusion, which prevents conjugation between donor cells containing Ti plasmids. Entry exclusion of the Ti plasmid is tightly regulated by TraR and mediated by TrbK

and TrbJ proteins. In the absence of OOHL, the Trb proteins are not expressed, and *Agrobacterium* cells harboring a Ti plasmid are efficient recipients. However, in the presence of OOHL, cells block the entry of Ti plasmids and instead become efficient conjugal donors.

Finally, the replication properties of the Ti plasmid were analyzed and I show that the origin of replication resides within the *repC* gene. I also show that RepC protein binds to a site located at an AT-rich region within its own gene. The DNA binding domain of RepC was localized to a region at the N-terminus of the protein. RepC functions *in cis* to initiate replication, and this mode of action may have important implications for plasmid compatibility.

BIOGRAPHICAL SKETCH

Uelinton Manoel Pinto was born on April 16, 1979 in Juiz de Fora, Minas Gerais, Brazil. He was raised in the rural surroundings of a small town called Chácara until he graduated from high school. He then moved to Viçosa to pursue his undergraduate studies at Federal University of Viçosa, receiving his degree in Food Engineering in 2003. In that same year he started his graduate studies in the department of Microbiology where he received a Master's degree in 2005, studying quorum sensing and the proteolytic activity of bacteria isolated from milk.

In 2005 he was awarded the prestigious CAPES/FULBRIGHT Fellowship from the Brazilian government to pursue his Ph.D. at Cornell University with Dr. Stephen C. Winans. In August of 2005 he came to Ithaca and has since worked with different aspects of the biology of the tumor inducing plasmid of *Agrobacterium tumefaciens.*

Para a minha família querida,

em especial minha mãe Marilda Duque Pinto

e meu pai João Alves Pinto (in memoriam)

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Stephen C. Winans for all the support, encouragement, advice, teachings, and availability to answer questions and discuss complex ideas. He has greatly influenced my scientific career, always prompting me to higher standards. Many thanks to my minor advisors Dr. John Helmann and Dr. Eric Alani for the encouragement, advice, and availability.

Special thanks go to the Department of Microbiology staff, specially Shirley Cramer, Cathy Shappell, Doreen Dineen, and Patti Butler for all the indispensible help they provide to us. I would like to thank my family back in Brazil for all their love and support over the years. I also wish to thank my colleagues from the Winans' lab, past and present members. Dr. Yungrong Chai and Dr. Catharine White for helping me to settle in the lab. Dr. Hongbaek Cho for passing along his expertise with experimental techniques. Dr. Anatol Eberhard for all the help with chemistry related tasks, and especially for bringing to the lab his humorous contribution. I thank the current lab members who have made this place such a nice working environment. Thanks to Esther D. Costa and Ana L. Flores Mireles for their friendship, encouragement, help in the lab, and discussions of scientific and non-scientific related topics. Thanks to Dr. Ching-Sung Tsai (Sonny) for friendship and help with techniques. Thanks to Gina Ryan for the friendship and good discussions. Thanks to Nidia Chen for the friendship. Thanks to Nathan Winans and Tom Winans for their help with lab supplies.

I would like to thank the Helmann's lab for "sharing" Dr. Ahmed and their equipment. Special thanks to Dr. Ahmed Gaballa and Dr. Shawn R. MacLellan for all the great tips with different techniques and for always being available. Thanks to my friends from the Helmann's Lab, specially Veronica Guariglia-Oropeza, Yun Luo, and Dr. Anna-Barbara Hachmann.

Thanks to the many friends I have made in Ithaca over the years. Their friendship was certainly very important for my success as a student and my personal growth. These include friends I made in the department, especially those I started the Ph.D. with. Thanks also to the many friends from all over the globe I made in the Cornell International Christian Fellowship (CICF). I also want to thank my church in Rochester, especially the Eppings (John, Elaine, and Julia) for their prayers, support and great care for me since I arrived in Ithaca.

I have also to mention Dr. Maria Cristina Dantas Vanneti who has greatly influenced my scientific career since my undergraduate studies. Thanks for all the encouragement, friendship, and advice.

Finally I would like to thank the Brazilian government for the financial support provided through a CAPES/FULBRIGHT scholarship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

Introduction

Agrobacterium tumefaciens belongs to the alpha-proteobacteria group and is part of the *Rhizobiaceae* family which includes plant pathogens and nitrogen-fixing symbionts (Slater *et al.*, 2009; Wood *et al.*, 2001). The bacterium is ubiquitous in soil and causes crown gall disease in dicotyledonous plants at wounded sites (Winans, 1992). The disease is characterized by the growth of tumors; it is usually not fatal, but can significantly reduce crop productivity (Escobar and Dandekar, 2003).

The first observations of crown gall disease date to the nineteenth century, but it took over 50 years for a bacterium to be identified as the causal agent of disease (Escobar and Dandekar, 2003; Smith and Townsend, 1907). Almost 70 years later the tumor inducing principle was linked to the presence of a large plasmid in the bacterial cells (Zaenen *et al.*, 1974). In a landmark study just a few years later, the tumor inducing principle was identified as a piece of DNA that was transferred from the bacterial plasmid to plant cells (Chilton *et al.*, 1977). Progress in research with what are now called tumor inducing (Ti) plasmids has moved quickly from that point on and has profoundly impacted many different fields including plant biology, agriculture, biotechnology and molecular biology (Binns, 2002; Escobar and Dandekar, 2003).

The Ti-plasmid is a large circular replicon of about 200 kb and carries the transferred DNA (also known as transforming or T-DNA) and most genes required for tumorigenesis (Pappas, 2008; White and Winans, 2007; Zhu *et al.*, 2000). The T-DNA carries a set of genes responsible for plant cell proliferation and another set of genes required for the synthesis of opines which support bacterial growth (Zhu *et al.*, 2000). The Ti-plasmid also codes for the transport and catabolism of opines produced in the

tumors. In fact, Ti-plasmids are usually classified according to the type of opines that are encoded in the T-DNA.

Infection starts when *A. tumefaciens* cells containing the Ti plasmid encounter a plant wounded site, which releases compounds such as amino acids, organic acis and sugars that activate the transfer of the T-DNA from the bacterial cells to plant cells. The T-DNA is transported as a complex with Vir proteins, some of which contain nuclear localicazion sequences that direct the T-DNA to the nucleus of the plant cell. Once the T-DNA is transported to the nucleus, it can integrate into the plant genome through a process that is not well understood and iniate expression of the tumor inducing and opine synthase genes (Escobar and Dandekar, 2003; Pappas, 2008; Zhu *et al.*, 2000).

Many aspects of Ti-plasmid biology have been extensively studied for many years. For instance, *A. tumefaciens* has become a model organism for studies on host-microbe interactions, inter-kingdom gene transfer, type IV secretion systems, and cell-to-cell communication commonly referred to as quorum sensing (Escobar and Dandekar, 2003; Farrand *et al.*, 2002; Zhu *et al.*, 2000). It is quite striking that some basic aspects of the Ti-plasmid biology are still poorly understood. Among them replication, segregation, and entry exclusion are topics with very limited information.

In this introductory chapter I will briefly describe the quorum sensing system in bacteria with focus on *A. tumefaciens,* and then I will review some aspects of plasmid entry exclusion, which is a phenomenon that prevents conjugation between cells containing the same plasmid. Finally, I will present a more comprehensive review about plasmid replication and segregation with special attention on the *repABC* type plasmids.

1.1. Quorum sensing in *Agrobacterium tumefaciens*

Quorum sensing is a phenomenon by which bacteria coordinate gene expression in response to small signaling molecules, also known as autoinducers, that accumulate as a function of cell density (Waters and Bassler, 2005; Whitehead *et al.*, 2001). The type of signal molecules can vary depending on the bacterial species and in that sense, quorum sensing systems are quite diverse. In Gram-negative bacteria, signaling is commonly mediated by acylhomoserine lactones (AHLs) while small peptides generally mediate cell-to-cell communication in Gram-positive cells (Platt and Fuqua, 2010; Waters and Bassler, 2005; Whitehead *et al.*, 2001). A group of small molecules termed autoinducer-2 (AI-2) is believed to control quorum sensing in both types of bacteria due to the widespread presence of its synthase LuxS (Fuqua and Greenberg, 2002). However, there is some controversy on whether AI-2 function is related to quorum sensing in species other than *vibrios* since the enzyme plays an important role in the activated methyl cycle (Rezzonico and Duffy, 2008).

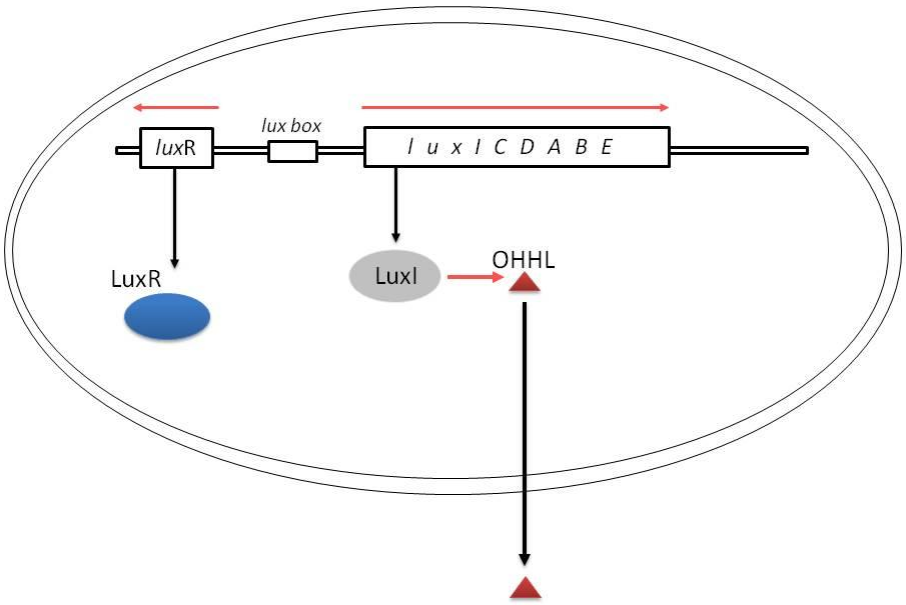
The paradigm for the AHL mediated quorum sensing system is found in the marine bacterium *Vibrio fischeri* (Nealson and Hastings, 1979). In this organism, quorum sensing relies on the production of N-3-oxohexanoyl-L-homoserinelatone (OHHL) by a protein called LuxI (Eberhard *et al.*, 1981). OHHL can diffuse through the bacterial cell membrane, accumulating in the light organ of certain species of squid and fish. When its concentration reaches a critical threshold, which coincides with high cell density, the molecule is able to diffuse back into the bacterial cell cytoplasm where it binds to a receptor protein called LuxR. LuxR bound to OHHL is a dimeric transcriptional activator that induces bioluminescence in *V. fischeri* (Fig. 1.1) (Antunes *et al.*, 2007; Pappas *et al.*, 2004; Waters and Bassler, 2005; Whitehead *et al.*, 2001). LuxR/LuxI homologous systems have been found to regulate diverse phenotypes in many proteobacteria, coordinating group behavior that ranges from

pathogenesis, biofilms formation, conjugation, production of secondary metabolites, to symbiosis (Waters and Bassler, 2005; Whitehead *et al.*, 2001). LuxR homologues are usually activated by a cognate AHL, however in a few cases these proteins are actually inactivated by their signal (Tsai and Winans, 2010). Another variation to the theme is the presence of LuxR homologues that are not coupled to a LuxI partner. These proteins, usually referred to as orphan regulators, are thought to either work in a ligand-independent fashion or to recognize endogenous or exogenous AHLs, as well as to detect other types of small molecules (Fuqua, 2006; Patankar and Gonzalez, 2009; Subramoni and Venturi, 2009). They can also function as dominant-negative inhibitors (Chai *et al.*, 2001).

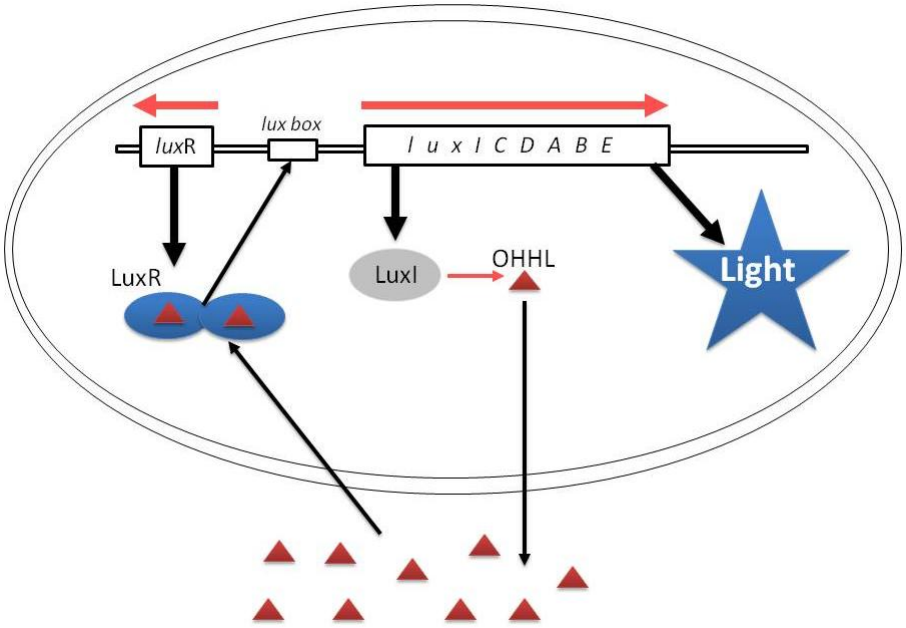
The crystal structure of TraR-OOHL complex bound to the *tra* box DNA has been independently solved by two groups (Vannini *et al.*, 2002; Zhang *et al.*, 2002). The structure confirms biochemical data showing that both the N-terminal and C-terminal domains contribute to protein dimerization (Qin *et al.*, 2000; Zhu and Winans, 2001). The main dimerization determinant is located in the N-terminal domain through the long hydrophobic α-helix 9, and the second dimerization element is at the C-terminal domain through α-helix 13 (Fig. 1.2). The OOHL molecule is fully buried inside a hydrophobic pocket located in the N-terminal domain, and makes no contact with the solvent (Zhang *et al.*, 2002).

Figure 1.1. The LuxR/I quorum sensing system of *Vibrio fisheri*.
A- At low cell densities LuxI synthesizes OHHL at basal levels which diffuses away from the cell.
B- At high cell densities, OOHL accumulates to a threshold concentration which allows it to diffuse back into the cell cytoplasm where it binds to LuxR protein. LuxR-OHHL complexes activate transcription of the *lux* operon producing bioluminescence. Red triangle denotes the autoincer OHHL.
C- Structure of OHHL produced by LuxI.

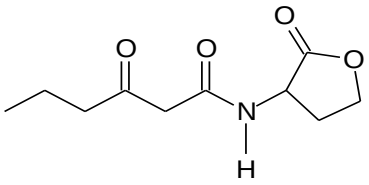
A


lux box
luxR
l u x I C D A B E
LuxR
LuxI
OHHL

B


lux box
luxR
l u x I C D A B E
LuxR
LuxI
OHHL
Light

C


O
O
O
O
N
H

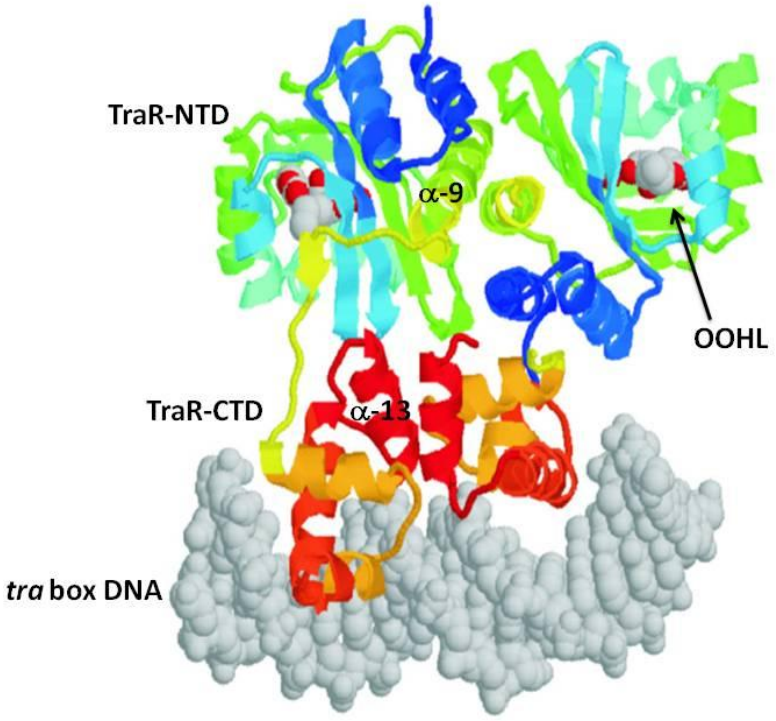


Figure 1.2. Crystal structure of TraR complexed with OOHL and *tra* box DNA. Dimerization helixes α9 and α13 are indicated for the left monomer. Each monomer binds one molecule of OOHL which is fully buried in the N-terminal domain (NTD) of the protein. Arrow points to the OOHL molecule. TraR-NTD is responsible for OOHL binding, dimerization and RNA polymerase (RNAP) contact. TraR-CTD binds DNA, aids in dimerization and provides further contact with RNAP. Structure coordinates can be found in (Zhang *et al.*, 2002).

Several studies have further confirmed and extended the structural predictions, broadening the understanding of TraR transcription activation and making it a pivotal model for the LuxR family of transcriptional regulators. For instance, site-directed mutagenesis studies have defined the OOHL interactions to the hydrophobic pocket of the protein (Chai and Winans, 2004; Luo *et al.*, 2003). The role of each base on the consensus *tra* box has also been evaluated. Some of the predicted interactions with the protein seem to be more important than others, while several non-contacted bases are crucial for TraR-DNA affinity, suggesting a role for DNA bending or flexibility (White and Winans, 2007). The role of dimerization has also been evaluated for TraR stability. Dimerization increases resistance to cellular proteases and together with OOHL binding, further contributes to TraR overall stability and function (Pinto and Winans, 2009; Zhu and Winans, 2001).

Other studies have also pointed out that specific residues of TraR might participate in RNA polymerase (RNAP) interaction. For instance, TraR must directly interact with RNAP, as transcription activation *in vitro* was shown to require only TraR-OOHL complexes, RNAP from *A. tumefaciens*, and promoter DNA containing the *tra* box (Zhu and Winans, 1999). In fact, several residues from both domains of TraR may be involved in the interaction with RNAP. Mutation analyses have unveiled an extensive array of contacts between TraR and RNAP at different classes of promoters, with distinct patches of the TraR protein participating in interactions with alpha and sigma subunits (Costa *et al.*, 2009; Luo and Farrand, 1999; Luo *et al.*, 2003; Qin *et al.*, 2004; Qin *et al.*, 2009; White and Winans, 2005). Further studies are required to determine the specific residues on the RNAP subunits that directly associate with TraR and to show how these interactions influence the transcription initiation and elongation processes.

1.2. Plasmid entry-exclusion

Bacterial conjugation, also referred to as bacterial mating or bacterial sex, is the process of plasmid or integrative conjugative element (ICE) transfer from a donor to a recipient cell. Conjugation plays a prominent role in horizontal gene transfer, disseminating a vast array of genetic information among microorganisms. Conjugative elements that are self-transmissible normally contain an origin of transfer, a relaxase, a coupling protein, a type IV secretion system, and an entry exclusion function (Marrero and Waldor, 2007a; Smillie *et al.*, 2010). The exclusion function of plasmids is believed to inhibit redundant conjugation between cells with similar or identical plasmids (Garcillan-Barcia and de la Cruz, 2008) and differs from the plasmid incompatibility mechanism which operates at the level of replication and/or segregation (Thomas and Smith, 1987). An entry exclusion function may be advantageous for a plasmid for three main reasons as pointed out by (Garcillan-Barcia and de la Cruz, 2008). First, it may avoid the entrance of an incompatible plasmid in the recipient cell; second, it may benefit the donor cell by avoiding unnecessary and energetically costly DNA transfer; and third, it may render the recipient cell immune to lethal zygosis which is a phenomenon caused by membrane damage due to excessive DNA conjugation (Ou, 1980).

Two different types of exclusion determinants are known to cause entry exclusion. Surface exposed outer membrane proteins like TraT have only been described for F-like plasmids (Garcillan-Barcia and de la Cruz, 2008), and are thought to block the formation of stable mating aggregates between two donor cells (Sukupolvi and O'Connor, 1990). Other conjugative plasmids have proteins such as TraS of the F plasmid and TrbK of RP4, which are located in the inner membrane and inhibit conjugative DNA transfer in a process that occurs downstream of mating pair formation (Garcillan-Barcia and de la Cruz, 2008; Haase *et al.*, 1996).

The interacting partners of TraS of F plasmids and the ICEs R391 and SXT have been described (Audette *et al.*, 2007; Marrero and Waldor, 2007a, b). Entry exclusion for these elements is mediated by interactions between the mating pair stabilization protein TraG in the donor and TraS in the recipient cells. The mechanism by which these proteins interact seems to be difficult to envision since both of them are located in the cytoplasmic membrane. (Marrero and Waldor, 2007a) suggested a model in which these interacting proteins must adopt a complex conformational change in order to interact and block DNA transfer. It is not known if the interaction happens in the donor or in the recipient cell, or if one of the proteins is translocated. No interacting partner of entry exclusion proteins has been characterized for any other plasmids besides F and the ICEs described above (Garcillan-Barcia and de la Cruz, 2008).

The Ti plasmids of *A. tumefaciens* are distinctive in the sense that they have a suite of genes for two types of conjugation systems. One that mediates T-DNA transfer to plant cells and another that is involved in Ti-plasmid transfer between bacterial cells. Entry exclusion does not apply for inter kingdom T-DNA transfer to plant cells (Garcillan-Barcia and de la Cruz, 2008). For the bacterium to bacterium conjugation the scenario might be different. (Hooykaas *et al.*, 1980) suggested that Ti plasmids do not encode for an entry exclusion function based on results showing that recipient cells containing a Ti plasmid were as good as recipient cells lacking it in conjugation assays. However, by homology search and sequencing analysis, several authors did not discard the possibility that Ti plasmids have such a system (Garcillan-Barcia and de la Cruz, 2008; Haase *et al.*, 1996; Li *et al.*, 1999). In fact, we have shown that two proteins, TrbK and TrbJ, of the Ti plasmid mediate robust entry exclusion and that these two proteins work synergistically (Cho *et al.*, 2009).

In general, the *trb* genes of the Ti plasmid share high similarity to the *tra2* region of IncP plasmids (exemplified by RP4, RK2, and R18) (Alt-Morbe *et al.*, 1996). For instance the TrbK protein has been shown to mediate entry exclusion of IncP type plasmids (Haase *et al.*, 1995; Haase *et al.*, 1996; Lessl *et al.*, 1991; Lyras *et al.*, 1994). The role of TrbJ of RP4 in entry exclusion has been a subject of debate. Two different groups have reported that TrbJ from IncPα plasmids mediated entry exclusion (Lessl *et al.*, 1991; Lyras *et al.*, 1994), while a third group found that TrbJ from RP4 plays no role in this process (Haase *et al.*, 1995; Haase *et al.*, 1996). The reasons for these conflicting data are unclear. TrbK proteins of RP4 and pTiC58 are not required for conjugation, and their sole function may be in entry exclusion. On the other hand the TrbJ proteins of both plasmids are essential for conjugation (Haase *et al.*, 1995; Li *et al.*, 1999).

TrbK of RP4 is a lipoprotein that has a lipid attachment motif and is localized mainly to the cytoplasmic membrane (Haase *et al.*, 1996). Its signal sequence is removed proteolytically and one or more acyl groups are added to a cysteine residue at the newly created amino terminus. This cysteine is required for wild type levels of entry exclusion, although residual levels were detectable when this cysteine was altered (Haase *et al.*, 1996). Alteration of the cysteine residue caused a decreased affinity for the cytoplasmic membrane. Significantly, all known Ti plasmid TrbK proteins lack this cysteine residue and are therefore unlikely to be acylated. It is not known whether the signal sequence of these proteins is removed, or where the mature proteins are localized. It is very likely that they localize to the cytoplasmic membrane.

All Ti-plasmid *tra* and *trb* genes are regulated by the TraR/TraI quorum-sensing system (Fuqua and Winans, 1994) and a variety of plasmids of *Rhizobium*, *Mesorhizobium*, and *Sinorhizobium* spp. regulate conjugation genes using a similar mechanism (Gonzalez and Marketon, 2003). Cell-to-cell communication

controls conjugation of the so-called pheromone plasmids of enterococci. The model plasmid for this system is pCF10 of *Enterococcus faecalis*. Transfer of pCF10 is induced in donor cells by small peptides produced by potential recipients (Chandler and Dunny, 2004; Dunny, 2007). The plasmid also codes for a surface exclusion protein that blocks conjugation between induced donor cells but does not interfere with conjugation to uninduced cells, even though they too contain a plasmid of the same type (Dunny *et al.*, 1985). Significantly, both TraR and TraI are encoded on Ti plasmids, and therefore this system seems to detect a quorum of conjugal donors rather than of conjugal recipients, as in the case of pCF10. The entry exclusion proteins TrbK and TrbJ are encoded in the *trb* operon, which is activated by quorum sensing. Therefore, this system presents some similarities to pCF10 in the sense that activated donor cells avoid conjugation among themselves and in the absence of quorum-sensing signals, these cells become efficient recipients (Fig. 1.3).

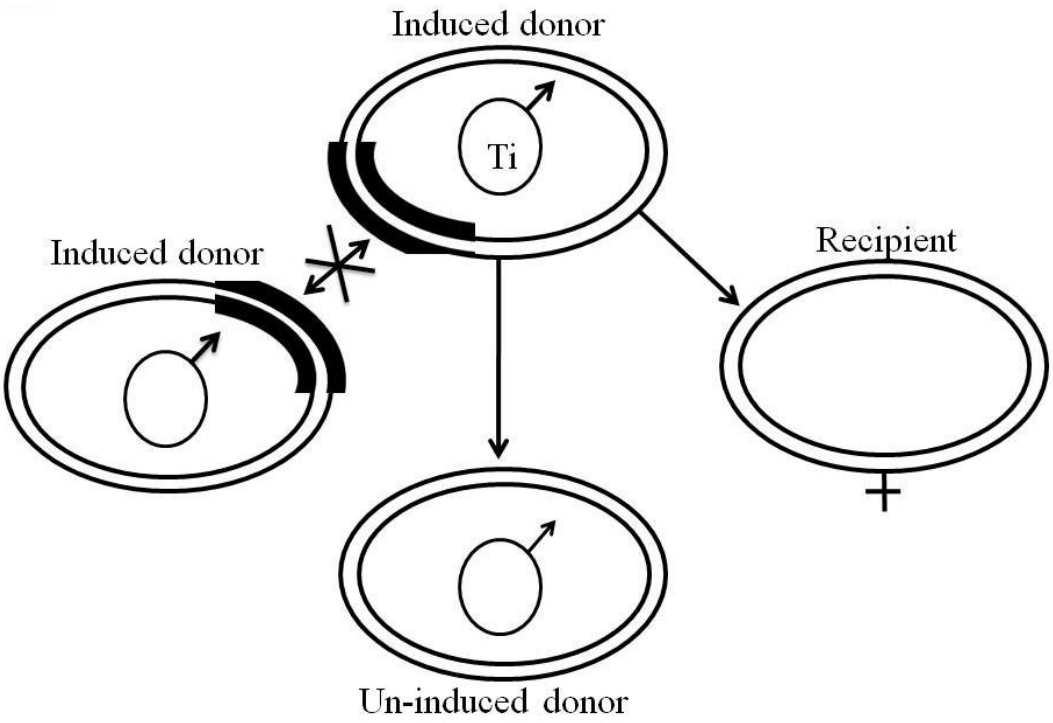


Figure 1.3. Entry exclusion in *A. tumefaciens.* Quorum sensing induced cells are unable to conjugate Ti plasmid among themselves (left) due to expression of entry exclusion proteins TrbK and TrbJ. Uninduced cells containing a Ti-plasmid (bottom) behave like recipients cells lacking a Ti-plasmid (right), and therefore are good recipients for conjugation.

1.3. The ABC of plasmid replication and segregation[1]

During each bacterial cell cycle, each replicon must ensure its efficient transmission to the daughter cells. Although plasmids can fail to do so, resulting in progeny cells that are cured of that plasmid, virtually all naturally occurring plasmids have evolved mechanisms to avoid their loss from the population. Ideally, plasmids must replicate on average once per cell cycle and have evolved both positive and negative replication factors to achieve this balance. Plasmids also generally encode mechanisms to physically segregate newly replicated copies into both daughter cells. I will apply these general principles to members of the RepABC family of bacterial plasmids, which includes the Ti plasmid of *A. tumefaciens,* which are ubiquitous throughout the alpha-proteobacteria group. RepA, RepB, and a partitioning site act in plasmid segregation to daughter cells, while RepC and the origin of replication are essential for plasmid replication initiation. In the majority of cases, these three genes are expressed as an operon, which may be regulated by environmental stimuli, resulting in fluctuations in gene dosage, and altered expression of every plasmid-encoded gene.

1.3.1. Replication and segregation in bacterial plasmids

Plasmids by definition are not essential for survival of the host. However, they frequently endow the cell with a complement of genes that impart extremely useful new physiologies, including the catabolism of new nutrients, the ability to cause diseases, to defend its host against antimicrobials, or to engage in mutualistic symbioses with host plants or animals. Plasmids can in some cases carry a large proportion of the genes in the bacterial genome (del Solar *et al.*, 1998). In many cases, these genes can be transmitted horizontally into new host bacteria by

---

[1] Part of a manuscript to be submitted for publication as a minireview in *Plasmid*.

conjugative transfer and can fundamentally alter the physiology of the recipient just minutes after the transfer event, promoting adaptation to diverse environmental challenges.

Plasmids have a variety of mechanisms that maximize their faithful vertical transmission from one mother cell to both daughter cells. First, their replication must occur on average once per cell cycle, so that sufficient copies exist to populate daughter cells. Second, plasmid segregation systems physically push or pull the plasmids into daughter cells (Thomas, 2000). Third, site-specific recombination systems convert plasmid multimers to monomers, which are simpler to partition to daughter cells. Fourth, postsegregational killing mechanisms cause plasmid-free cells to lose viability, while plasmid-containing cells are protected by an "antidote" protein (Engelberg-Kulka and Glaser, 1999). Fifth, as described above, conjugative transfer in some cases provides a mechanism for plasmid-free cells to reacquire the same plasmid by horizontal transfer. In addition, conjugative plasmids generally block the entry of identical or similar plasmids into the host cell, a phenomenon known as entry exclusion (Cho *et al.*, 2009). This characteristic is believed to be essential for plasmid survival, first because all conjugative plasmids code for at least one entry exclusion gene, and second because it may free a plasmid from competition during segregation into daughter cells (Garcillan-Barcia and de la Cruz, 2008; Kues and Stahl, 1989).

The strong selection for stable transmission of plasmids has caused these systems to have evolved several times independently. Several families of replication initiator proteins have been studied, with no apparent evolutionary similarity (del Solar *et al.*, 1998). At least two families of partitioning systems (Gerdes *et al.*, 2000) and many different post-segregational killing systems have evolved (Van Melderen and Saavedra De Bast, 2009). In contrast, it is probable that conjugative transfer

evolved just once, although these systems have undergone such divergent evolution that they lose obvious signs of homology.

The RepABC family of replication initiators and partitioning systems is widespread throughout the alpha-proteobacteria group (Castillo-Ramirez *et al.*, 2009; Cevallos *et al.*, 2008; Petersen *et al.*, 2009; Slater *et al.*, 2009). The conserved function of these genes, their organization as an operon, the interactions between RepA and RepB proteins, and their overall gene arrangement argue for a monophyletic origin. However, various subfamilies have been described, indicating a role for lateral gene transfer events of individual genes within the operon, which may increase the chances of plasmid compatibility (Castillo-Ramirez *et al.*, 2009; Cevallos *et al.*, 2002; Cevallos *et al.*, 2008; Crossman *et al.*, 2008; Petersen *et al.*, 2009; Slater *et al.*, 2009). In fact, a study in the *Rhodobacterales*, a group within the alpha-proteobacteria, revealed the presence of nine compatibility groups based on phylogenetic analysis of *repABC* genes, further arguing for the diversity within this replicon type**.** These systems appear to vastly outnumber all other mechanisms for replication and partitioning in the alpha-proteobacteria. Besides numerous plasmids, the secondary chromosomes of *Agrobacterium* and *Brucella* species are also replicated by a *repABC*-type replicon (Paulsen *et al.*, 2002; Slater *et al.*, 2009). In addition, proteins homologous to RepA and RepB, involved in partitioning of plasmids and chromosomes, are found within and beyond the alpha group (Petersen *et al.*, 2009).

Simple homology searches can be useful in examining the diversity of RepABC proteins. Using BLASTP to find proteins homologous to RepC of the *Agrobacterium* octopine type Ti plasmid, I found 218 apparent homologues, every one of which was a member of the alpha group. In contrast, RepAB homologues are found throughout a much larger group of bacteria. The *repABC* operon structure appears to be highly conserved, although a few plasmids of the alpha group contain *repAB*

homologues while lacking a *repC* gene (Castillo-Ramirez *et al.*, 2009). Interestingly, a plasmid from the *Rhodobacterales* contains a *repAB* cassette, but the replicator initiator protein is a new type that shares homology to DnaA (Petersen *et al.*, 2010). Conversely, plasmid pTiBo542 of *A. tumefaciens*, the so-called supervirulent Ti plasmid, contains two tandem *repABC* operons. It is not known whether both are functional. Additionally, some plasmids contain a *repC* gene but don't have an associated *repAB* counterpart (Bartosik *et al.*, 1997; Burgos *et al.*, 1996; Castillo-Ramirez *et al.*, 2009; Izquierdo *et al.*, 2005).

Of the hundreds of members of the *repABC* family, none has received nearly as much attention as analogous systems in other plasmids. The best studied plasmid replication and partitioning systems are those of the F plasmid, the P1 prophage, plasmid R6K, the high copy number plasmid ColE1, the broad host range plasmid RK2, and the gram positive plasmid pT181 of *Staphylococcus aureus*. Among the *repABC* group, a few have begun to be studied beyond the bioinformatic level. These include that of the symbiosis plasmids p42d of *Rhizobium etli* and pSymA of *Sinorhizobium meliloti*, two tumor inducing (Ti) plasmids of *A. tumefaciens*, and plasmid pTAV1 of *Paracoccus versutus* (Bartosik *et al.*, 1997; Bartosik *et al.*, 1998; Cevallos *et al.*, 2002; Cevallos *et al.*, 2008; Chai and Winans, 2005a; MacLellan *et al.*, 2005; Pappas, 2008; Venkova-Canova *et al.*, 2004). This review will focus primarily on biochemical and genetic studies of the small number of *repABC* systems that have been so far characterized. Bioinformatic insights gleaned from the hundreds of other such systems will be included as appropriate. Phylogenetic studies of this group have been described elsewhere (Bartosik *et al.*, 2002b; Castillo-Ramirez *et al.*, 2009; Palmer *et al.*, 2000; Petersen *et al.*, 2009; Turner *et al.*, 1996).

The genomic architecture of the four bacteria above mentioned is quite dissimilar. *R. etli* CFN42 has one circular chromosome and six circular plasmids, all

having *repABC* cassettes (Gonzalez *et al.*, 2006). Of these p42d carries the complement of *nod, nif,* and *fix* genes required for plant nodulation, and has therefore received the most attention. *S. meliloti* 1021 contains one circular chromosome and two mega plasmids (1.35 and 1.7 megabases) (Galibert *et al.*, 2001). Plasmid pSymA has the majority of *nod, nif,* and *fix* genes and is curable, while pSymB encodes many genes for synthesis of an exopolysaccharide that is required for nodulation. It also encodes one essential gene (a tRNA), so technically it should be considered a chromosome rather than a plasmid. pSymA and pSymB both have *repABC* cassettes. *A. tumefaciens* C58 has a circular chromosome (2.84 MB), a linear chromosome (2.1 MB), the Ti plasmid (0.21 MB), which contains most of the genes required for pathogenesis, and a second plasmid pAT (0.54 MB) (Goodner *et al.*, 2001; Wood *et al.*, 2001). The linear chromosome and both plasmids replicate via *repABC* cassettes. *P. versutus* has not been sequenced in its entirety, but appears to have one circular chromosome and two plasmids, pTAV1 and pTAV3 (Dolowy *et al.*, 2005). Of these, pTAV1 has two replication origins, one of which has a complete *repABC* cassette, while the other seems more diverse containing only the *repC* gene (Bartosik *et al.*, 1997). The other plasmid, pTAV3, does not have a *repABC* cassette (Bartosik *et al.*, 2002a).

1.3.2. Structure of *repABC* type cassettes

Most plasmids contain replication and partitioning genes at different sites in the plasmid genomes and therefore regulate these functions independently. The *repABC*-type plasmids are an exception as those genes tend to be organized in an operon structure. Proper expression of these operons is essential for plasmid survival as both replication and partitioning can be affected concomitantly. We shall see that some of these plasmids have elaborated regulatory mechanisms such as quorum

sensing and autorepression which fine-tune the expression of the operon according to the environment conditions.

Genetic maps of several *repABC* cassettes are shown in Fig. 1.4. Four of the five have just one of each gene, while pTAV1 has an extra copy of *repC* that is fully functional for replication but not for stable partitioning (Bartosik *et al.*, 1998). Despite the apparent gaps between some of these genes, there is no evidence for internal promoters. The available data indicates that all three genes are expressed solely from promoters upstream of *repA*. In pTiA6, which is considered to be functionally identical to pTiR10, mutations that block activity of promoters upstream of *repA* block *repC* function. We are not aware of similar work in other members of this family. The gaps between these genes turn out to contain interesting features. In the case of pTiA6, there is a gene, *repD*, 76 codons in length, that fully spans the gap between *repA* and *repB*. A *repD-lacZ* fusion indicates that *repD* is translated (Chai and Winans, 2005b). Plasmid pTiC58 has a similar reading frame that is identical in length, but completely divergent in sequence. This divergence suggests that the RepD protein may have no function whatsoever. It is believed that translation of this gene may be required to prevent the appearance of an untranslated mRNA between *repA* and *repB*, which could cause transcription termination (Chai and Winans, 2005b). Within *repD* of pTiA6 are two perfect direct repeats that are highly conserved in pTiC58 and some other members of this gene family that appear to serve as partitioning sites. The biological role of these sites in pTiC58 has not been investigated. In contrast, no small gene is found between *repA* and *repB* of p42d, pSymA, or pTAV1. It turns out that partition sites in these plasmids lie outside of the *repABC* operon (Fig. 1.4).

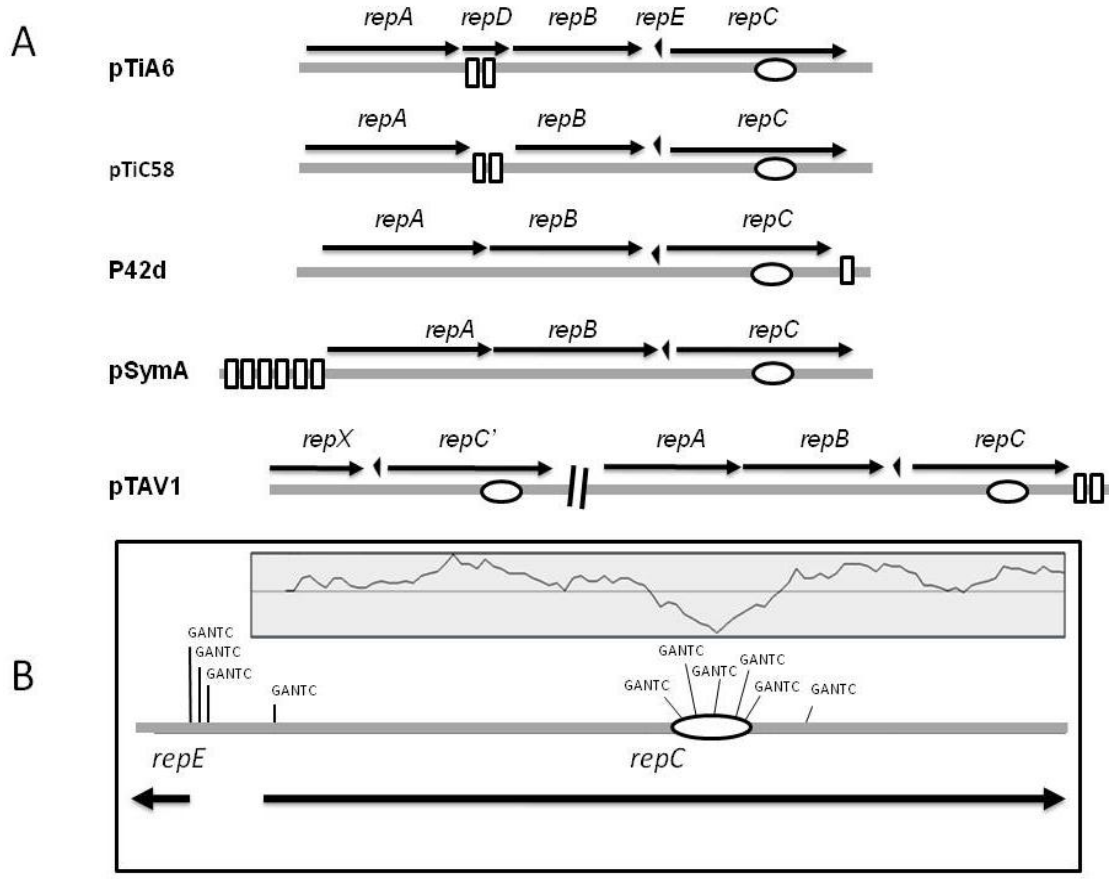


Figure 1.4. Genetic organization of *repABC* operons from representative plasmids. (A) *repABC* module of *A. tumefaciens* pTiA6 and pTiC58, from *R. etli* CFN42 plasmid p42d, from *S. meliloti* plasmid pSymA, and from *P. versutus* pTAV1 (see text for details). Rectangles represent the partitioning sites, white ovals the AT-rich region believed to be *oriV*. Inverted arrow head represents the small antisense RNA, which in the case of pTiA6 is called *repE*.
(B) *repE-repC* region of the Ti plasmid showing the abundance of GANTC sites located on the *repE* promoter and on the AT-rich region. The GC content is shown in the box above *repC* gene.

` A similar gap is seen between *repB* and *repC* of all five plasmids, and is also found in all members of the *repABC* gene family. This region has long been known to confer incompatibility between like plasmids (Tabata *et al.*, 1989), but only recently it was found to encode an antisense RNA molecule that functions to downregulate expression of *repC* and mediate copy number control (Chai and Winans, 2005a; MacLellan *et al.*, 2005; Venkova-Canova *et al.*, 2004). There is no evidence for a *repC* promoter in this region. All expression of these genes appears to initiate upstream of *repA*, and changes in promoter expression can have profound effects on RepC function and consequent plasmid copy number.

There are two other notable features of these operons. First, each has a conspicuous AT-rich region within *repC* (white ovals in Fig. 1.4). As described later, this region is thought to be the origin of plasmid replication. Second, the nucleotide sequence GANTC is overrepresented in the putative replication origin and in the promoter of the antisense RNA (Fig. 1.4.B). These sequences are substrates for a methylase found in *Caulobacter crescentus, A. tumefaciens,* and most alpha-proteobacteria (Brilli *et al.*, 2010). These enzymes methylate the A residue of GANTC on each strand, and are thought to play a central role in the cell cycle of these organisms.

1.3.3. Plasmid partitioning by RepA and RepB

RepA and RepB belong to the broad family of ParA/ParB-type partitioning proteins, used by many plasmids, phages and chromosomes in widely different bacteria (Bignell and Thomas, 2001; Ebersbach and Gerdes, 2005; Gerdes *et al.*, 2000). These systems also require a *cis*-acting site, often referred to as *parS* or *parC*, which are functionally analogous to the centromeres of eukaryotic chromosomes.

These sites are usually, but not always, found closely linked to the *parAB* genes (Breier and Grossman, 2007; Grigoriev and Lobocka, 2001).

The RepA and RepB proteins of *repABC* systems are members of the Type 1a family of partitioning systems. Most members of this subfamily are encoded by bi-cistronic operons with the ParA-type protein promoter proximal. Like other ParA homologs in this group, RepA proteins contain amino-terminal helix-turn-helix motifs and carboxy-terminal Walker-type ATPases that energize the partitioning process (Gerdes *et al.*, 2000). Other members of this group are found in very low-copy plasmids and prophages P1 and P7. The best studied members of this clade are SopA of the F plasmid and ParA of plasmid P1 (Davey and Funnell, 1994; Davis *et al.*, 1992; Mori *et al.*, 1989). The low copy number of these replicons underscores the need for an accurate partitioning system. Several ParA-type proteins form dynamic intracellular filaments that affect DNA positioning in the predivisional cell or DNA translocation during segregation (Barilla *et al.*, 2005; Derman *et al.*, 2008; Ebersbach and Gerdes, 2004; Hatano *et al.*, 2007; Ringgaard *et al.*, 2009). ParA undergoes a slow and specific conformational change upon ATP binding which enables it to bind non-specific DNA such as the nucleoid (Vecchiarelli *et al.*, 2010). In fact, ParA dynamically moves along the nucleoid, and by interacting with the partitioning complex composed of ParB and the *cis*-acting *parC* DNA site, it is able to position plasmids to daughter cells (Gerdes *et al.*, 2010; Howard and Gerdes, 2010; Ringgaard *et al.*, 2009). The mechanism of positioning is not yet clear but some models predict the nucleoid working as a matrix where ParA-ATP builds up in concentration attracting the ParB-plasmid complex. Then ParB stimulates ParA ATPase activity with subsequent disassembly from the nucleoid, and the plasmid is moved either by the pulling force generated by the depolymerization of ParA or by the pulling force directed towards regions of high ParA-ATP concentration (Ringgaard *et al.*, 2009;

Vecchiarelli *et al.*, 2010). In addition, the DNA binding domain of ParA proteins facilitates specific binding to an operator site near the promoter which causes negative autoregulation of transcription in a process that does not require ATP (Dunham *et al.*, 2009; Vecchiarelli *et al.*, 2010).

RepB proteins resemble SopB of F and ParB of P1 (Davis and Austin, 1988; Lobocka and Yarmolinsky, 1996; Surtees and Funnell, 2001; Watanabe *et al.*, 1989). Members of the ParB family, also called centromere-binding proteins, are structurally diverse, multi-domain proteins and can be grouped as those containing a helix-turn-helix DNA binding domain or those containing a ribbon-helix-helix DNA fold (Schumacher, 2007, 2008). These proteins directly contact *parS* sites via one of the above mentioned DNA motifs and interact with ParA usually via the N-terminal end, stimulating the ATPase and repressor activities of their partners (Adachi *et al.*, 2006; Bignell and Thomas, 2001; Dye and Shapiro, 2007; Gerdes *et al.*, 2010; Lee and Grossman, 2006; Schumacher, 2007). Some ParB-type proteins bind DNA up to several kilobases flanking the *parS* sites, and in so doing, can interfere with expression of genes in this region (Bingle *et al.*, 2005; Breier and Grossman, 2007; Rodionov *et al.*, 1999).

A number of genetic studies of partitioning of RepABC-type plasmids have been reported including those of pTiB6SE (similar to other octopine-type Ti plasmids such as pTiR10, and pTiA6), pTAV320, p42d, and pSymA. Genetic and molecular analyses of these plasmids showed that insertions, frame-shift mutations, or deletions in *repA* or *repB* substantially decrease plasmid stability (Bartosik *et al.*, 1998; Gallie *et al.*, 1985; MacLellan *et al.*, 2006; Ramirez-Romero *et al.*, 2000; Tabata *et al.*, 1989).

The *parS* loci of pTAV320, p42d, pSymA, pTiC58 and pTiR10 have been identified. They consist in one or more copies of a 16-bp palindromic consensus sequence (GTTNNCNGCNGNNAAC) and fulfill three requisites: first, they are

essential for plasmid stability; second, they are RepB-binding sites; and finally, they are incompatible with their respective parental plasmid when provided *in trans* (Bartosik *et al.*, 1998; Chai and Winans, 2005b; MacLellan *et al.*, 2006; Venkova-Canova *et al.*, 2004). The number and position of par-site elements present in this family of replicons vary widely (Fig. 1.4). Plasmid p42d contains just one such site, located just beyond the 3' end of *repC* (Venkova-Canova *et al.*, 2004). Plasmid pTAV1 contains two such sites, also located just downstream of *repC* (Bartosik et al., 2001). Plasmid pTiA6 also contains two sites, but these are located between *repA* and *repB* (Chai and Winans, 2005b). Specifically, they lie within the *repD* minigene, and deletion of one of these two sites did not seem to impair partitioning (Chai and Winans, 2005b). Plasmid pSymA has six such elements, located upstream of the *repABC* promoter (MacLellan *et al.*, 2006). Cloning a partitioning site from pSymA onto a compatible plasmid has resulted in incompatibility between the two replicons. This is presumably due to competition between the two plasmids for partitioning machinery, which apparently is limiting in concentration. Point mutations in the cloned partitioning site eliminated plasmid instability and reduced affinity for *repB* protein (MacLellan *et al.*, 2006).

The positioning of *repABC* replicon origins of *A. tumefaciens* and *S. meliloti* has been studied using fluorescence in situ hybridization (FISH). The origin region of all replicons of both species localized to or near the cell pole, while the origin of the broad host range plasmid RK2 did not. In double labeling experiments in *A. tumefaciens*, it was shown that two origin regions of two different replicons rarely colocalized; rather, they occupied nearby but clearly distinct sites at the pole (Kahng and Shapiro, 2003). It will be of great interest to study the location and movement of the origins and the partition complexes of the *repABC* replicons in live cells to see if they follow a similar path as described for ParA/ParB mediated segregation in other

low copy number plasmids (Ringgaard *et al.*, 2009; Vecchiarelli *et al.*, 2010). It is hard to envision a nucleoid assisted movement of *repABC* plasmids and secondary chromosomes considering that some of these elements have sizes comparable to the main chromosome. Another level of complication comes from the fact that some bacteria have multiple *repABC* type replicons. It is likely that a novel mechanism is responsible for partitioning in these systems.

1.3.4. RepC and the origin of replication.

Unlike RepA and RepB, RepC proteins have been identified only within the alpha-proteobacteria (Bartosik *et al.*, 1997; Burgos *et al.*, 1996; Castillo-Ramirez *et al.*, 2009; Palmer *et al.*, 2000; Petersen *et al.*, 2009). RepC proteins bear no significant homologies to any other replication initiator proteins, and so far, no biochemical analysis of a RepC protein has been published. This review must therefore focus on the available genetic studies. Several RepC proteins have been shown to be essential and sufficient for plasmid replication (Izquierdo *et al.*, 2005; Ramirez-Romero *et al.*, 2000; Tabata *et al.*, 1989). Narrow host range plasmids into which *repC* has been cloned gained the ability to replicate in the bacterium whence that *repC* gene originated (Bartosik *et al.*, 1997; Chai and Winans, 2005a; Izquierdo *et al.*, 2005; Mercado-Blanco and Olivares, 1994). Most *repC* genes are located within *repABC* cassettes, though a few exceptions have been reported (Fig. 1.4) (Bartosik *et al.*, 1997; Izquierdo *et al.*, 2005; Mercado-Blanco and Olivares, 1994).

A few studies showing that *repC* suffices for replication initiation also suggest that the origin of replication lies within the *repC* gene (Bartosik *et al.*, 1998; Cevallos *et al.*, 2008). Most studies of the type have shown that a region upstream of *repC* (located between *repB* and *repC*) was also required for replication (Chai and Winans, 2005a; Izquierdo *et al.*, 2005; MacLellan *et al.*, 2005; Venkova-Canova *et al.*, 2004).

The role of this region has been implicated in the control of *repC* expression as we shall see below. The localization of a replication origin within the coding sequence of the replication initiator has been reported for plasmids pAD1 of *Enterococcus faecalis*, pSX267 of *Staphylococcus xylosus*, pSK41 of *S. aureus*, pLS32 of *Bacillus subtilis*, and the N15 prophage of *E. coli.* It is also found on the lytic origin of bacteriophage lambda. The origins of these and many other plasmids contain so-called iterons, or directly repeated DNA sequences (Francia *et al.*, 2004; Gering *et al.*, 1996; Kwong *et al.*, 2004; Ravin *et al.*, 2003; Tanaka *et al.*, 2005). *repABC* origins lack any apparent iterons. However, the central portion of *repC* genes contains an AT-rich sequence of about 150 nucleotides, which is another common feature of replication origins. Based on this assumption, several authors have suggested that the origin of replication resides within this region (Bartosik *et al.*, 1998; Cevallos *et al.*, 2008; Petersen *et al.*, 2009; Wagner-Dobler *et al.*, 2010). In fact, I was able to show that purified RepC from the *A. tumefaciens* octopine-type Ti plasmid binds to a region located within the AT-rich segment and that mutations in this site, that conserved the amino acid sequence, abolished replication (Pinto and Winans, manuscript in preparation). Cloning of *repC* under controlled expression shows that the upstream region is not required for autonomous replication, and its role may solely be related to downregulation of *repC* expression and plasmid compatibility (see below). Secondary structure predictions and amino acid conservation predict that RepC may have two domains, an N-terminal domain (NTD) from residues 1-242 connected to the C-terminal domain (CTD) from residues 297-439 by an unstructured linker (Fig. 1.5). A fragment containing just residues 26-158 of RepC is sufficient for binding to the putative *oriV*, though at somewhat lower affinity than wild type (Pinto and Winans, manuscript in preparation). This region of the NTD has a structural resemblance to the MarR family of transcription factors which bind DNA by means of a winged helix-

turn-helix motif. Overexpression of RepC increases copy number of the plasmid expressing it but does not affect the copy number of another *repC* plasmid, indicating that the protein works *in cis* (Pinto and Winans, manuscript in preparation). Like other replication initiator proteins from iteron plasmids, RepC does not have an ATPase domain (Chattoraj, 2000). The role of the RepC C-terminal domain remains to be revealed.

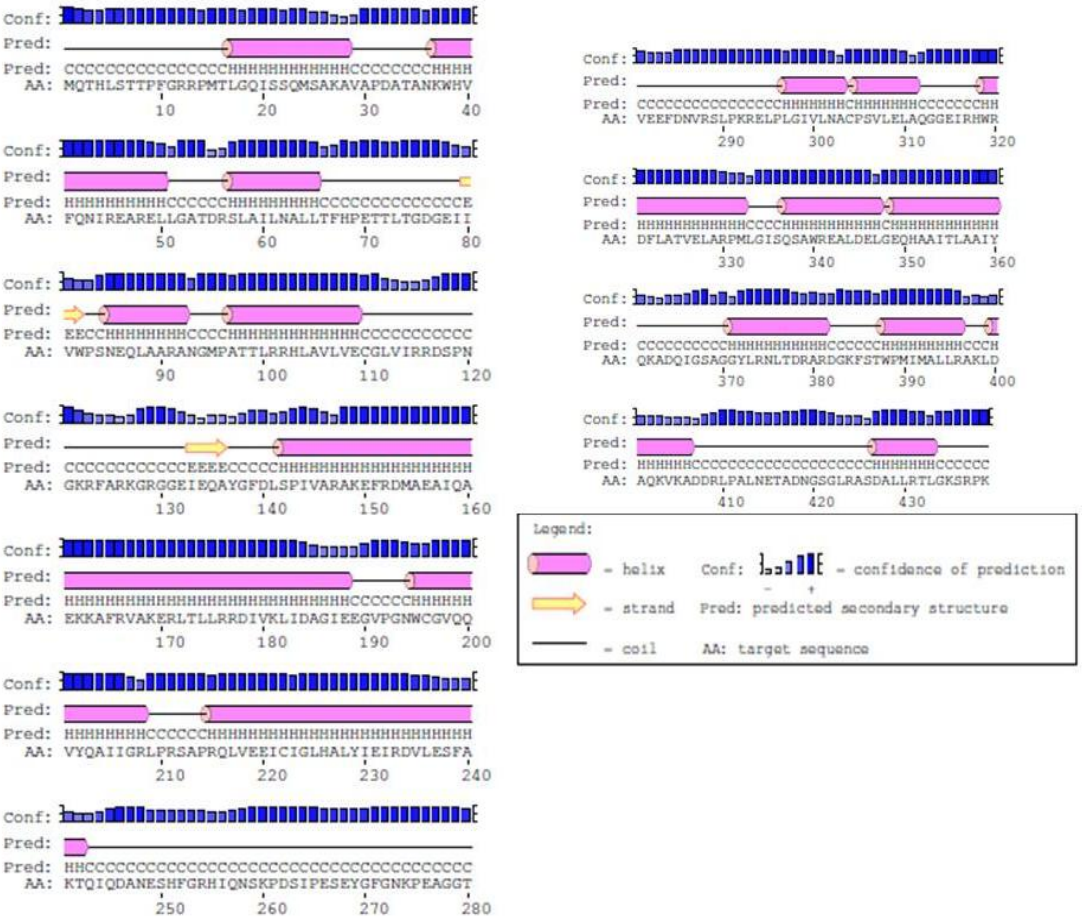


Figure 1.5. Secondary structure prediction of RepC (PSIPRED - http://bioinf.cs.ucl.ac.uk/psipred/). The NTD, comprised of residues 1-242, is responsible for DNA binding and presents structural homology to the MarR family of transcriptional factors. The CTD has no known function and is linked to NTD through a linker from residues 243 to 296.

Replication initiator proteins (Rep) from the iteron-type plasmids generally bind to the origin of replication in a monomeric state and recruit DnaA. Binding to the origin is usually cooperative and these interactions are believed to induce DNA melting at the AT-rich region, creating the replicative bubble, similar to the chromosomal initiation events at *oriC* in *E. coli*. The Rep-DnaA-DNA complex recruits the helicase (DnaB in *E. coli*) and replication usually follows the θ-type (del Solar et al., 1998). There are no iterative sequences and no DnaA binding sites matching the consensus for the alpha-proteobacteria group anywhere within *repC* (Brilli et al., 2010). It nonetheless seems plausible that RepC could recruit the replicative helicase with the help of DnaA, as it has been shown that DnaA can bind to more relaxed consensus sequences (Messer, 2002). It is also possible that RepC is able to induce strand separation and recruit the helicase protein on its own.

Plasmids must use host encoded proteins in order to replicate. It is quite puzzling that several *repABC* replicons, including the Ti plasmid, carry a copy of the *dnaE* gene, which codes for a homolog of the α subunit of the DNA polymerase III (Slater et al., 2009; Wood et al., 2001). The maintenance of mini-*repC* replicons does not require this gene, as it has been shown in other studies (Chai and Winans, 2005a; Li and Farrand, 2000; Pappas and Winans, 2003a). However, *dnaE* may have some secondary role in plasmid replication, helping to lower the burden to the cells in order to replicate the multipartite genomes of *Agrobacterium* species. This idea needs to be experimentally addressed in future studies.

1.3.5. Regulation of transcription initiation affects copy number

As described above, all transcription of *repA -B,* and *-C* genes probably initiates upstream of *repA.* Information about transcriptional regulation is sparse for most of these systems, but has been well developed for the pTiR10. The *repABC*

operon of this plasmid is transcribed from no fewer than four promoters (Fig. 1.6). Promoters P1, P2, and P3 are activated by the quorum-sensing transcription factor TraR, which requires the pheromone 3-oxooctanoylhomoserine lactone (OOHL) for activity. These promoters are transcribed divergently from the *traI* promoter, and expression of these divergent promoters requires overlapping TraR binding sites. Specifically, expression of the *traI* promoter requires a site called *tra* box II, which is also required for activation of promoters P1 and P2. This *tra* box is centered 42.5 nucleotides upstream of promoter P1 and 62.5 nucleotides upstream of promoter P2. Another TraR binding site, *tra* box III, is centered 42.5 nucleotides upstream of P3. Therefore, promoters P1 and P3 resemble those of class II promoters, while P2 resembles those of class I promoters as described for CRP (Busby and Ebright, 1999). The activation of these promoters by TraR leads to a moderate increase in Ti plasmid copy number (Pappas and Winans, 2003a). Copy number of pTiC58 is also enhanced by TraR-OOHL complexes, although the mechanisms have not yet been worked out (Li and Farrand, 2000). Similar data have also been described for the symbiotic plasmid pRL1JI of *Rhizobium leguminosarum* (Danino *et al.*, 2003). In contrast, plasmid p42d has just one promoter expressing its *repABC* operon, and this promoter is expressed constitutively (Ramirez-Romero *et al.*, 2001).

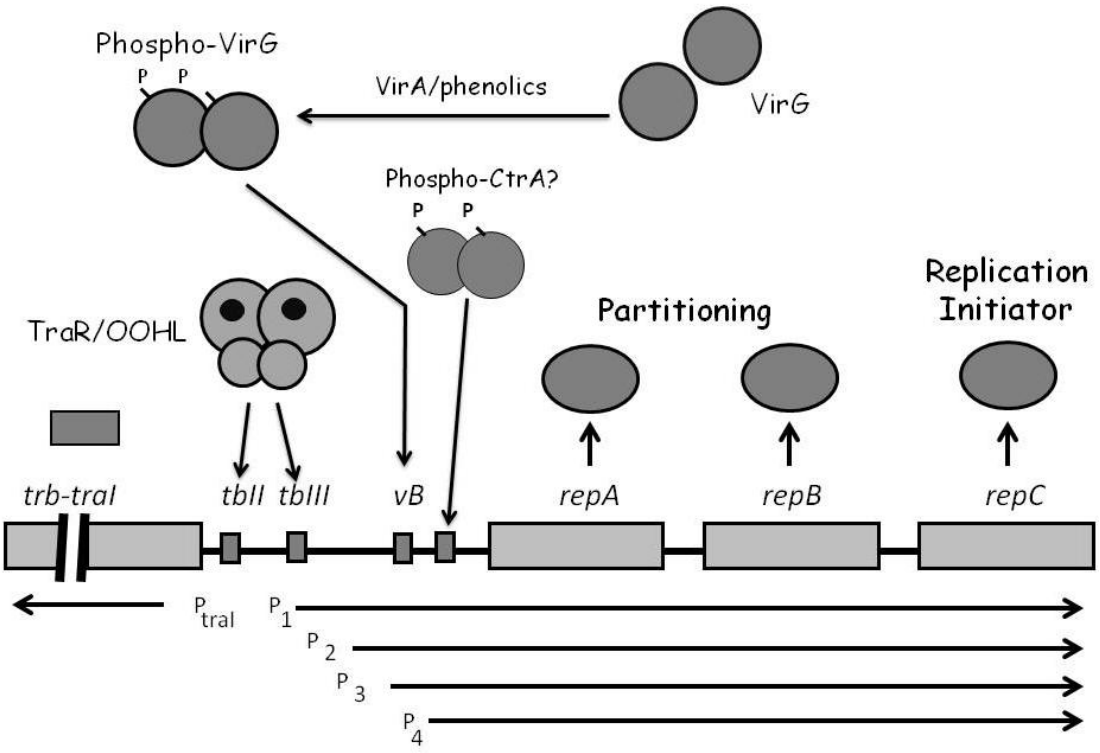


Figure 1.6. Positive regulation of the *repABC* operon of the octopine-Ti plasmid. TraR/OOHL complexes bind to *tra* box (tb) II and III to activate transcription of *traI-trb* operon, and of the *repABC* operon through promoters P1-P3. VirG, when phosphorylated by VirA, activates transcription of promoter P4. Additional regulation may be provided by Phospho-CtrA.

Promoter P4 provides basal level expression of the operon and is also activated by the two component response regulator VirG (Fig. 1.6) (Cho and Winans, 2005). Therefore, expression of the *repABC* operon of the octopine-type Ti plasmid is enhanced both by plant-released chemical signals (which lead to phosphorylation of VirG), and by quorum sensing pheromones (which activate TraR). In both cases, increased *repABC* expression leads to increased Ti plasmid copy number. The P4 promoter region also contains a possible binding site for another two-component response regulator, CtrA, which regulates the cell cycle of *Caulobacter crescentus* (Fig. 1.6). This binding site is centered around 50 nucleotides upstream of promoter P4. The significance of this motif, if any, remains to be explored.

Transcription initiation of the *repABC* operon of pTiR10 is also subjected to negative autoregulation (Fig. 1.7). The RepA protein binds to a site directly downstream of promoter P4, and strongly represses promoter expression (Pappas and Winans, 2003b). RepB alone does not bind this region and does not repress expression of P4; however, RepB enhances the ability of RepA to do so. This is probably due to the formation of a RepA-RepB complex that binds the P4 operator more strongly than does RepA alone. The converse is also true. RepB alone can bind to the two partitioning sites that lie between *repA* and *repB*. Binding to these sites blocks RNA polymerase read-through. RepA enhances the ability of RepB to bind these sites and inhibit RNA polymerase (Chai and Winans, 2005b). Similar data have been described for p42d (Ramirez-Romero *et al.*, 2001). For this plasmid, the partitioning site enhances repression, despite the fact that it is located at the 3' end of *repC* (Soberon *et al.*, 2004). A DNA loop model was proposed consisting of RepA-RepB complexes binding to the *repABC* promoter via RepA, and binding simultaneously to the partitioning site via RepB (Chai and Winans, 2005b).

Further transcription repression of this operon may come from binding of RepC protein to o*riV* located within the coding region of *repC* gene. This binding would work as a transcription road block, preventing the passage of RNA polymerase through this site similarly to what happens with RepB at the partitioning sites. This idea remains to be tested.

1.3.6. Regulation by antisense RNA

All of the *repABC* operons have an antisense RNA lying between *repB* and *repC* (Fig. 1.4). These antisense RNA molecules are approximately 50 nucleotides in length, and have a predicted stem loop structure followed by a run of U residues (Fig. 1.8). These sequences may serve as rho-independent transcription terminators, and also may play a role in contacting complementary sequences on the RepABC mRNA (Cervantes-Rivera *et al.*, 2010; Chai and Winans, 2005a). The promoters of these antisense RNA molecules have consensus σ-70 motifs, and fusions between these genes and *lacZ* synthesize high levels of β-galactosidase. Similar RNA molecules are predicted to be conserved in many or perhaps all members of the *repABC* operon family. Antisense RNAs are also found in other types of plasmid replication systems, and may have evolved independently (Weaver, 2007).

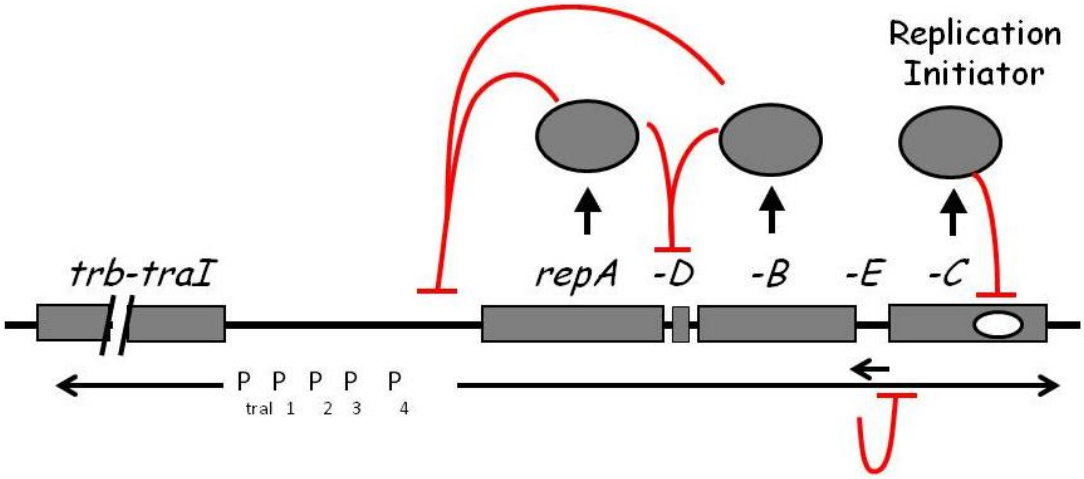


Figure 1.7. Negative auto-regulation of the *repABC* operon of the octopine-Ti plasmid. Transcription of the *repABC* operon is inhibited by autorepression mediated by RepA/RepB complexes at the operator region downstream of P4 and at the partitioning site located within *repD*. *repC* expression is inhibited by the small antisense RNA RepE which downregulates transcription and translation of *repC*. Further inhibition may be mediated by autorepression when RepC binds to *oriV*.

The antisense RNA of pTiR10 is termed RepE. RepE inhibits both the translation and the continued transcription of *repC* (Chai and Winans, 2005a). Point mutations in the RepE promoter that decrease its activity lead to increases in plasmid copy number. RepE is predicted to alter the balance between two alternative stem loop structures in the mRNA upstream of *repC*. According to this model, in the absence of RepE, the ribosome binding site is available for translation initiation, while in the presence of RepE, an alternative stem loop sequesters the ribosome binding site, and could also serve as a transcription terminator (Fig. 1.8). In fact, this model was confirmed for plasmid p42d by Cervantes-Rivera and collaborators (2010). These authors observed that the target RNA (the mRNA upstream of *repC*) adopts two different conformations depending on the presence or absence of the small antisense RNA. They confirmed that in the absence of the antisense RNA, the target mRNA folds in a way that the Shine-Dalgarno sequence is free for translation. However, when the antisense RNA pairs with the target, it changes the conformation of the mRNA forming two stem-loop structures one of which operates as a transcription terminator to abort *repC* transcription, meanwhile obstructing the ribosomal binding site of the remaining transcripts that were fully synthesized (Cervantes-Rivera *et al.*, 2010). Therefore, expression of *repC* in a wild type operon, under non-inducing conditions, can be so low that *repC-lacZ* translational fusions do not synthesize detectable levels of β-galactosidase (Pappas and Winans, 2003b).

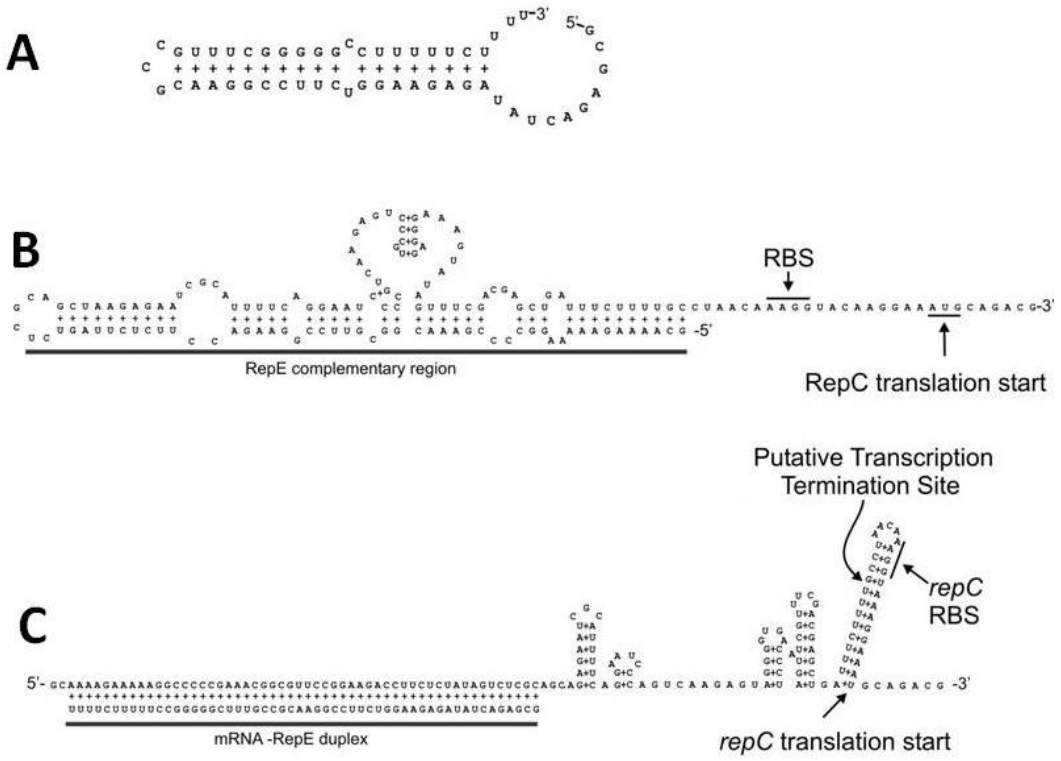


Figure 1.8. Model of the mechanism of action of *repE* on the control of *repC* expression.
(A) Predicted structure of RepE.
(B) Predicted structure of the mRNA upstream of *repC* in the absence of RepE. Note that the ribosomal binding site (RBS) and the translation starting site are available for translation initiation.
(C) Binding of RepE to the target mRNA upstream of *repC* causes a conformational change that induces formation of several hairpins, one of which resembles a Rho-independent transcription terminator. *repC* expression is further down-regulated by sequestration of the RBS and the translation starting site in the terminator hairpin. These figures were adapted from (Chai and Winans 2005a).

Plasmids expressing just the antisense RNA gene are incompatible with the parent plasmid, indicating that the RNA is a trans-acting incompatibility determinant. Point mutations that alter the structure or that decrease the expression of this RNA decrease incompatibility (Cervantes-Rivera *et al.*, 2010; Chai and Winans, 2005a; Li and Farrand, 2000; MacLellan *et al.*, 2005; Ramirez-Romero *et al.*, 2000; Tabata *et al.*, 1989; Venkova-Canova *et al.*, 2004).

1.3.7. Regulation by DNA methylation

DNA methylases play many important roles in bacterial physiology (Casadesus and Low, 2006). Perhaps the best known class is the one that protects DNA from cognate restriction endonucleases. Other DNA methylases have no role in protection against restriction endonucleases. Instead, these methylases, exemplified by Dam methylase of *E. coli*, are believed to distinguish parental DNA from newly synthesized daughter DNA, which is transiently unmethylated. Distinguishing between parental and daughter DNA is important in mismatch repair, initiation of chromosome replication, and in the expression of some genes (Marinus and Casadesus, 2009). The *E. coli* origin of replication is rich in GATC sites, which are methylated at the N6 position of A residues on both strands. After replication, hemimethylated GATC sites are tightly bound by the SeqA protein, which transiently protects these sites from methylation. SeqA also sequesters the origin from the DnaA protein, which is required for the next round of replication. Hemimethylated GATC sites therefore delay the next round of replication (Katayama *et al.*, 2010).

In the alpha-proteobacterium *Caulobacter crescentus*, there is an analogous (though not homologous) methylase called CcrM, which methylates the N6 position of A residues in the sequence GANTC (Wion and Casadesus, 2006). For example, the origin of replication has five GANTC sites, and methylation of those is required for

activation of the *Caulobacter* origin of replication (Shaheen et al., 2009). GANTC sites are also found in the promoters of several genes that play important roles in the timing of the cell cycle (Collier *et al.*, 2007). It has been shown that methylation of these promoters regulate gene expression (Collier *et al.*, 2007; Reisenauer and Shapiro, 2002). *C. crescentus* does not have a protein homologous to SeqA, but some other protein could play an analogous role.

There is one important difference between Dam and CcrM. The former is thought to be active at all points of the cell cycle, while the latter is synthesized only by predivisional cells. Therefore, GATC sites of *E. coli* are hemimethylated only transiently, while hemimethylated GANTC sites are methylated only at a particular point in the *C. crescentus* cell cycle (Casadesus and Low, 2006; Collier *et al.*, 2007).

The majority of alpha-proteobacteria genomes have homologues of CcrM that could play important roles in the cell cycle of these organisms (Brilli *et al.*, 2010). We have purified the CcrM protein of *A. tumefaciens* and confirmed that it methylates GANTC sites in *vitro* (unpublished data). Significantly, the putative origins of replication of all *repABC* plasmids are rich in GANTC sites. Similarly, the promoter of the antisense RNA of these plasmids is also rich in these sites (Fig. 1.4.B). If methylation of these sequences occurs only at a particular point in the cell cycle, this could have important consequences for origin utilization and/or expression of the RepC proteins. These ideas remain to be tested.

Expression of the Ti plasmid *rep* operon in *A. tumefaciens* combines a sophisticated array of gene regulation. As I have described, this complex network is comprised of both positive regulators, in the form of environmental stimuli (plant released signals) and quorum sensing, and negative regulators, in the form of autorepression and a small antisense RNA. The possible control through methylation

of GANTC sites and CtrA binding to the *repABC* promoter sequence may further synchronize replication of the Ti plasmid with the *A. tumefaciens* cell cycle.

1.4. Scope of the dissertation

This work concentrates on three aspects of the Ti-plasmid biology. The first study deals with quorum sensing, and in experiments done with the LuxR-type transcriptional regulator called TraR, I evaluated the contribution of dimerization to overall TraR stability and function. The second study, which was done in collaboration with Dr. Hongbaek Cho, shows that entry exclusion of the Ti-plasmid is mediated by TrbJ and TrbK and is regulated by quorum sensing. The third and last part of the dissertation deals with vegetative replication of the Ti-plasmid and the essential involvement of *repC* in this process.

REFERENCES

Adachi, S., Hori, K., and Hiraga, S. (2006) Subcellular positioning of F plasmid mediated by dynamic localization of SopA and SopB. *J Mol Biol* **356**: 850-863.

Alt-Morbe, J., Stryker, J.L., Fuqua, C., Li, P.L., Farrand, S.K., and Winans, S.C. (1996) The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. *J Bacteriol* **178**: 4248-4257.

Antunes, L.C., Schaefer, A.L., Ferreira, R.B., Qin, N., Stevens, A.M., Ruby, E.G., and Greenberg, E.P. (2007) Transcriptome analysis of the *Vibrio fischeri* LuxR-LuxI regulon. *J Bacteriol* **189**: 8387-8391.

Audette, G.F., Manchak, J., Beatty, P., Klimke, W.A., and Frost, L.S. (2007) Entry exclusion in F-like plasmids requires intact TraG in the donor that recognizes its cognate TraS in the recipient. *Microbiology* **153**: 442-451.

Barilla, D., Rosenberg, M.F., Nobbmann, U., and Hayes, F. (2005) Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *Embo J* **24**: 1453-1464.

Bartosik, D., Wlodarczyk, M., and Thomas, C.M. (1997) Complete nucleotide sequence of the replicator region of *Paracoccus* (*Thiobacillus*) *versutus* pTAV1 plasmid and its correlation to several plasmids of *Agrobacterium* and *Rhizobium* species. *Plasmid* **38**: 53-59.

Bartosik, D., Baj, J., and Wlodarczyk, M. (1998) Molecular and functional analysis of pTAV320, a *repABC*-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. *Microbiology* **144 (Pt 11)**: 3149-3157.

Bartosik, D., Szymanik, M., and Wysocka, E. (2001) Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J Bacteriol* **183**: 6234-6243.

Bartosik, D., Baj, J., Bartosik, A.A., and Wlodarczyk, M. (2002a) Characterization of the replicator region of megaplasmid pTAV3 of *Paracoccus versutus* and search for plasmid-encoded traits. *Microbiology* **148**: 871-881.

Bartosik, D., Baj, J., Piechucka, E., Waker, E., and Wlodarczyk, M. (2002b) Comparative characterization of *repABC*-type replicons of *Paracoccus pantotrophus* composite plasmids. *Plasmid* **48**: 130-141.

Bignell, C., and Thomas, C.M. (2001) The bacterial ParA-ParB partitioning proteins. *J Biotechnol* **91**: 1-34.

Bingle, L.E., Macartney, D.P., Fantozzi, A., Manzoor, S.E., and Thomas, C.M. (2005) Flexibility in repression and cooperativity by KorB of broad host range IncP-1 plasmid RK2. *J Mol Biol* **349**: 302-316.

Binns, A.N. (2002) T-DNA of *Agrobacterium tumefaciens*: 25 years and counting. *Trends Plant Sci* **7**: 231-233.

Breier, A.M., and Grossman, A.D. (2007) Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol Microbiol* **64**: 703-718.

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., and Biondi, E.G. (2010) The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. *BMC Syst Biol* **4**: 52.

Burgos, P.A., Velazquez, E., and Toro, N. (1996) Identification and distribution of plasmid-type A replicator region in rhizobia. *Mol Plant Microbe Interact* **9**: 843-849.

Busby, S., and Ebright, R.H. (1999) Transcription activation by catabolite activator protein (CAP). *J Mol Biol* **293**: 199-213.

Casadesus, J., and Low, D. (2006) Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* **70**: 830-856.

Castillo-Ramirez, S., Vazquez-Castellanos, J.F., Gonzalez, V., and Cevallos, M.A. (2009) Horizontal gene transfer and diverse functional constrains within a common replication-partitioning system in Alphaproteobacteria: the *repABC* operon. *BMC Genomics* **10**: 536.

Cervantes-Rivera, R., Romero-Lopez, C., Berzal-Herranz, A., and Cevallos, M.A. (2010) Analysis of the mechanism of action of the antisense RNA that controls the replication of the *repABC* plasmid p42d. *J Bacteriol* **192**: 3268-3278.

Cevallos, M.A., Porta, H., Izquierdo, J., Tun-Garrido, C., Garcia-de-los-Santos, A., Davila, G., and Brom, S. (2002) *Rhizobium etli* CFN42 contains at least three plasmids of the *repABC* family: a structural and evolutionary analysis. *Plasmid* **48**: 104-116.

Cevallos, M.A., Cervantes-Rivera, R., and Gutierrez-Rios, R.M. (2008) The *repABC* plasmid family. *Plasmid* **60**: 19-37.

Chai, Y., Zhu, J., and Winans, S.C. (2001) TrlR, a defective TraR-like protein of *Agrobacterium tumefaciens*, blocks TraR function in vitro by forming inactive TrlR:TraR dimers. *Mol Microbiol* **40**: 414-421.

Chai, Y., and Winans, S.C. (2004) Site-directed mutagenesis of a LuxR-type quorum-sensing transcription factor: alteration of autoinducer specificity. *Mol Microbiol* **51**: 765-776.

Chai, Y., and Winans, S.C. (2005a) A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol Microbiol* **56**: 1574-1585.

Chai, Y., and Winans, S.C. (2005b) RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between repA and repB for plasmid partitioning and autorepression. *Mol Microbiol* **58**: 1114-1129.

Chandler, J.R., and Dunny, G.M. (2004) Enterococcal peptide sex pheromones: synthesis and control of biological activity. *Peptides* **25**: 1377-1388.

Chattoraj, D.K. (2000) Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol Microbiol* **37**: 467-476.

Chilton, M.D., Drummond, M.H., Merio, D.J., Sciaky, D., Montoya, A.L., Gordon, M.P., and Nester, E.W. (1977) Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. *Cell* **11**: 263-271.

Cho, H., and Winans, S.C. (2005) VirA and VirG activate the Ti plasmid *repABC* operon, elevating plasmid copy number in response to wound-released chemical signals. *Proc Natl Acad Sci U S A* **102**: 14843-14848.

Cho, H., Pinto, U.M., and Winans, S.C. (2009) Transsexuality in the rhizosphere: quorum sensing reversibly converts *Agrobacterium tumefaciens* from phenotypically female to male. *J Bacteriol* **191**: 3375-3383.

Collier, J., McAdams, H.H., and Shapiro, L. (2007) A DNA methylation ratchet governs progression through a bacterial cell cycle. *Proc Natl Acad Sci U S A* **104**: 17111-17116.

Costa, E.D., Cho, H., and Winans, S.C. (2009) Identification of amino acid residues of the pheromone-binding domain of the transcription factor TraR that are required for positive control. *Mol Microbiol* **73**: 341-351.

Crossman, L.C., Castillo-Ramirez, S., McAnnula, C., Lozano, L., Vernikos, G.S., Acosta, J.L., Ghazoui, Z.F., Hernandez-Gonzalez, I., Meakin, G., Walker, A.W., Hynes, M.F., Young, J.P., Downie, J.A., Romero, D., Johnston, A.W.,

Davila, G., Parkhill, J., and Gonzalez, V. (2008) A common genomic framework for a diverse assembly of plasmids in the symbiotic nitrogen fixing bacteria. *PLoS One* **3**: e2567.

Danino, V.E., Wilkinson, A., Edwards, A., and Downie, J.A. (2003) Recipient-induced transfer of the symbiotic plasmid pRL1JI in *Rhizobium leguminosarum* bv. *viciae* is regulated by a quorum-sensing relay. *Mol Microbiol* **50**: 511-525.

Davey, M.J., and Funnell, B.E. (1994) The P1 plasmid partition protein ParA. A role for ATP in site-specific DNA binding. *J Biol Chem* **269**: 29908-29913.

Davis, M.A., and Austin, S.J. (1988) Recognition of the P1 plasmid centromere analog involves binding of the ParB protein and is modified by a specific host factor. *Embo J* **7**: 1881-1888.

Davis, M.A., Martin, K.A., and Austin, S.J. (1992) Biochemical activities of the parA partition protein of the P1 plasmid. *Mol Microbiol* **6**: 1141-1147.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M., and Diaz-Orejas, R. (1998) Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**: 434-464.

Derman, A.I., Lim-Fong, G., and Pogliano, J. (2008) Intracellular mobility of plasmid DNA is limited by the ParA family of partitioning systems. *Mol Microbiol* **67**: 935-946.

Dolowy, P., Mondzelewski, J., Zawadzka, R., Baj, J., and Bartosik, D. (2005) Cloning and characterization of a region responsible for the maintenance of megaplasmid pTAV3 of *Paracoccus versutus* UW1. *Plasmid* **53**: 239-250.

Dunham, T.D., Xu, W., Funnell, B.E., and Schumacher, M.A. (2009) Structural basis for ADP-mediated transcriptional regulation by P1 and P7 ParA. *Embo J* **28**: 1792-1802.

Dunny, G.M., Zimmerman, D.L., and Tortorello, M.L. (1985) Induction of surface exclusion (entry exclusion) by *Streptococcus faecalis* sex pheromones: use of monoclonal antibodies to identify an inducible surface antigen involved in the exclusion process. *Proc Natl Acad Sci U S A* **82**: 8582-8586.

Dunny, G.M. (2007) The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philos Trans R Soc Lond B Biol Sci* **362**: 1185-1193.

Dye, N.A., and Shapiro, L. (2007) The push and pull of the bacterial cytoskeleton. *Trends Cell Biol* **17**: 239-245.

Eberhard, A., Burlingame, A.L., Eberhard, C., Kenyon, G.L., Nealson, K.H., and Oppenheimer, N.J. (1981) Structural identification of autoinducer of *Photobacterium fischeri* luciferase. *Biochemistry* **20**: 2444-2449.

Ebersbach, G., and Gerdes, K. (2004) Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Mol Microbiol* **52**: 385-398.

Ebersbach, G., and Gerdes, K. (2005) Plasmid segregation mechanisms. *Annu Rev Genet* **39**: 453-479.

Engelberg-Kulka, H., and Glaser, G. (1999) Addiction modules and programmed cell death and antideath in bacterial cultures. *Annu Rev Microbiol* **53**: 43-70.

Escobar, M.A., and Dandekar, A.M. (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci* **8**: 380-386.

Farrand, S.K., Qin, Y., and Oger, P. (2002) Quorum-sensing system of *Agrobacterium* plasmids: analysis and utility. *Methods Enzymol* **358**: 452-484.

Francia, M.V., Fujimoto, S., Tille, P., Weaver, K.E., and Clewell, D.B. (2004) Replication of *Enterococcus faecalis* pheromone-responding plasmid pAD1:

location of the minimal replicon and *oriV* site and RepA involvement in initiation of replication. *J Bacteriol* **186**: 5003-5016.

Fuqua, C., and Winans, S.C. (1996) Localization of OccR-activated and TraR-activated promoters that express two ABC-type permeases and the *traR* gene of Ti plasmid pTiR10. *Mol Microbiol* **20**: 1199-1210.

Fuqua, C., and Greenberg, E.P. (2002) Listening in on bacteria: acyl-homoserine lactone signalling. *Nat Rev Mol Cell Biol* **3**: 685-695.

Fuqua, C. (2006) The QscR quorum-sensing regulon of *Pseudomonas aeruginosa*: an orphan claims its identity. *J Bacteriol* **188**: 3169-3171.

Fuqua, W.C., and Winans, S.C. (1994) A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J Bacteriol* **176**: 2796-2806.

Galibert, F., Finan, T.M., Long, S.R., Puhler, A., Abola, P., Ampe, F., Barloy-Hubler, F., Barnett, M.J., Becker, A., Boistard, P., Bothe, G., Boutry, M., Bowser, L., Buhrmester, J., Cadieu, E., Capela, D., Chain, P., Cowie, A., Davis, R.W., Dreano, S., Federspiel, N.A., Fisher, R.F., Gloux, S., Godrie, T., Goffeau, A., Golding, B., Gouzy, J., Gurjal, M., Hernandez-Lucas, I., Hong, A., Huizar, L., Hyman, R.W., Jones, T., Kahn, D., Kahn, M.L., Kalman, S., Keating, D.H., Kiss, E., Komp, C., Lelaure, V., Masuy, D., Palm, C., Peck, M.C., Pohl, T.M., Portetelle, D., Purnelle, B., Ramsperger, U., Surzycki, R., Thebault, P., Vandenbol, M., Vorholter, F.J., Weidner, S., Wells, D.H., Wong, K., Yeh, K.C., and Batut, J. (2001) The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science* **293**: 668-672.

Gallie, D.R., Hagiya, M., and Kado, C.I. (1985) Analysis of *Agrobacterium tumefaciens* plasmid pTiC58 replication region with a novel high-copy-number derivative. *J Bacteriol* **161**: 1034-1041.

Garcillan-Barcia, M.P., and de la Cruz, F. (2008) Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid* **60**: 1-18.

Gerdes, K., Moller-Jensen, J., and Bugge Jensen, R. (2000) Plasmid and chromosome partitioning: surprises from phylogeny. *Mol Microbiol* **37**: 455-466.

Gerdes, K., Howard, M., and Szardenings, F. (2010) Pushing and pulling in prokaryotic DNA segregation. *Cell* **141**: 927-942.

Gering, M., Gotz, F., and Bruckner, R. (1996) Sequence and analysis of the replication region of the *Staphylococcus xylosus* plasmid pSX267. *Gene* **182**: 117-122.

Gonzalez, J.E., and Marketon, M.M. (2003) Quorum sensing in nitrogen-fixing rhizobia. *Microbiol Mol Biol Rev* **67**: 574-592.

Gonzalez, V., Santamaria, R.I., Bustos, P., Hernandez-Gonzalez, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S.C., Ramirez, M.A., Jimenez-Jacinto, V., Collado-Vides, J., and Davila, G. (2006) The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc Natl Acad Sci U S A* **103**: 3834-3839.

Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B.S., Cao, Y., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., Doughty, D., Scott, C., Lappas, C., Markelz, B., Flanagan, C., Crowell, C., Gurson, J., Lomo, C., Sear, C., Strub, G., Cielo, C., and Slater, S. (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294**: 2323-2328.

Grigoriev, P.S., and Lobocka, M.B. (2001) Determinants of segregational stability of the linear plasmid-prophage N15 of *Escherichia coli*. *Mol Microbiol* **42**: 355-368.

Haase, J., Lurz, R., Grahn, A.M., Bamford, D.H., and Lanka, E. (1995) Bacterial conjugation mediated by plasmid RP4: RSF1010 mobilization, donor-specific phage propagation, and pilus production require the same Tra2 core components of a proposed DNA transport complex. *J Bacteriol* **177**: 4779-4791.

Haase, J., Kalkum, M., and Lanka, E. (1996) TrbK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncP alpha plasmid RP4. *J Bacteriol* **178**: 6720-6729.

Hatano, T., Yamaichi, Y., and Niki, H. (2007) Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Mol Microbiol* **64**: 1198-1213.

Hooykaas, P.J., den Dulk-Ras, H., Ooms, G., and Schilperoort, R.A. (1980) Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. *J Bacteriol* **143**: 1295-1306.

Howard, M., and Gerdes, K. (2010) What is the mechanism of ParA-mediated DNA movement? *Mol Microbiol* **78**: 9-12.

Izquierdo, J., Venkova-Canova, T., Ramirez-Romero, M.A., Tellez-Sosa, J., Hernandez-Lucas, I., Sanjuan, J., and Cevallos, M.A. (2005) An antisense RNA plays a central role in the replication control of a *repC* plasmid. *Plasmid* **54**: 259-277.

Kahng, L.S., and Shapiro, L. (2003) Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J Bacteriol* **185**: 3384-3391.

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010) Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*. *Nat Rev Microbiol* **8**: 163-170.

Kues, U., and Stahl, U. (1989) Replication of plasmids in gram-negative bacteria. *Microbiol Rev* **53**: 491-516.

Kwong, S.M., Skurray, R.A., and Firth, N. (2004) Staphylococcus aureus multiresistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense RNA regulation. *Mol Microbiol* **51**: 497-509.

Lee, P.S., and Grossman, A.D. (2006) The chromosome partitioning proteins Soj (ParA) and Spo0J (ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in *Bacillus subtilis*. *Mol Microbiol* **60**: 853-869.

Lessl, M., Krishnapillai, V., and Schilf, W. (1991) Identification and characterization of two entry exclusion genes of the promiscuous IncP plasmid R18. *Mol Gen Genet* **227**: 120-126.

Li, P.L., Hwang, I., Miyagi, H., True, H., and Farrand, S.K. (1999) Essential components of the Ti plasmid *trb* system, a type IV macromolecular transporter. *J Bacteriol* **181**: 5033-5041.

Li, P.L., and Farrand, S.K. (2000) The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. *J Bacteriol* **182**: 179-188.

Lobocka, M., and Yarmolinsky, M. (1996) P1 plasmid partition: a mutational analysis of ParB. *J Mol Biol* **259**: 366-382.

Luo, Z.Q., and Farrand, S.K. (1999) Signal-dependent DNA binding and functional domains of the quorum-sensing activator TraR as identified by repressor activity. *Proc Natl Acad Sci U S A* **96**: 9009-9014.

Luo, Z.Q., Smyth, A.J., Gao, P., Qin, Y., and Farrand, S.K. (2003) Mutational analysis of TraR. Correlating function with molecular structure of a quorum-sensing transcriptional activator. *J Biol Chem* **278**: 13173-13182.

Lyras, D., Chan, A.W., McFarlane, J., and Stanisich, V.A. (1994) The surface exclusion system of RP1: investigation of the roles of *trbJ* and *trbK* in the surface exclusion, transfer, and slow-growth phenotypes. *Plasmid* **32**: 254-261.

MacLellan, S.R., Smallbone, L.A., Sibley, C.D., and Finan, T.M. (2005) The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. *Mol Microbiol* **55**: 611-623.

MacLellan, S.R., Zaheer, R., Sartor, A.L., MacLean, A.M., and Finan, T.M. (2006) Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons. *Mol Microbiol* **59**: 1559-1575.

Marinus, M.G., and Casadesus, J. (2009) Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol Rev* **33**: 488-503.

Marrero, J., and Waldor, M.K. (2007a) Determinants of entry exclusion within Eex and TraG are cytoplasmic. *J Bacteriol* **189**: 6469-6473.

Marrero, J., and Waldor, M.K. (2007b) The SXT/R391 family of integrative conjugative elements is composed of two exclusion groups. *J Bacteriol* **189**: 3302-3305.

Mercado-Blanco, J., and Olivares, J. (1994) The large nonsymbiotic plasmid pRmeGR4a of *Rhizobium meliloti* GR4 encodes a protein involved in replication that has homology with the RepC protein of *Agrobacterium* plasmids. *Plasmid* **32**: 75-79.

Messer, W. (2002) The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev* **26**: 355-374.

Mori, H., Mori, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A., and Hiraga, S. (1989) Purification and characterization of SopA and SopB proteins essential for F plasmid partitioning. *J Biol Chem* **264**: 15535-15541.

Nealson, K.H., and Hastings, J.W. (1979) Bacterial bioluminescence: its control and ecological significance. *Microbiol Rev* **43**: 496-518.

Ou, J.T. (1980) Role of surface exclusion genes in lethal zygosis in *Escherichia coli* K12 mating. *Mol Gen Genet* **178**: 573-581.

Palmer, K.M., Turner, S.L., and Young, J.P. (2000) Sequence diversity of the plasmid replication gene *repC* in the *Rhizobiaceae*. *Plasmid* **44**: 209-219.

Pappas, K.M., and Winans, S.C. (2003a) A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol Microbiol* **48**: 1059-1073.

Pappas, K.M., and Winans, S.C. (2003b) The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. *Mol Microbiol* **49**: 441-455.

Pappas, K.M., Weingart, C.L., and Winans, S.C. (2004) Chemical communication in proteobacteria: biochemical and structural studies of signal synthases and receptors required for intercellular signalling. *Mol Microbiol* **53**: 755-769.

Pappas, K.M. (2008) Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. *Plasmid* **60**: 89-107.

Patankar, A.V., and Gonzalez, J.E. (2009) Orphan LuxR regulators of quorum sensing. *FEMS Microbiol Rev* **33**: 739-756.

Paulsen, I.T., Seshadri, R., Nelson, K.E., Eisen, J.A., Heidelberg, J.F., Read, T.D., Dodson, R.J., Umayam, L., Brinkac, L.M., Beanan, M.J., Daugherty, S.C., Deboy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Nelson, W.C., Ayodeji, B., Kraul, M., Shetty, J., Malek, J., Van Aken, S.E., Riedmuller, S., Tettelin,

H., Gill, S.R., White, O., Salzberg, S.L., Hoover, D.L., Lindler, L.E., Halling, S.M., Boyle, S.M., and Fraser, C.M. (2002) The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proc Natl Acad Sci U S A* **99**: 13148-13153.

Petersen, J., Brinkmann, H., Berger, M., Brinkhoff, T., Pauker, O., and Pradella, S. Origin and evolution of a novel DnaA-like plasmid replication type in *Rhodobacterales*. *Mol Biol Evol*.

Petersen, J., Brinkmann, H., and Pradella, S. (2009) Diversity and evolution of *repABC* type plasmids in *Rhodobacterales*. *Environ Microbiol* **11**: 2627-2638.

Petersen, J., Brinkmann, H., Berger, M., Brinkhoff, T., Pauker, O., and Pradella, S. (2010) Origin and evolution of a novel DnaA-like plasmid replication type in *Rhodobacterales*. *Mol Biol Evol* **in press**.

Pinto, U.M., and Winans, S.C. (2009) Dimerization of the quorum-sensing transcription factor TraR enhances resistance to cytoplasmic proteolysis. *Mol Microbiol* **73**: 32-42.

Piper, K.R., Beck Von Bodman, S., Hwang, I., and Farrand, S.K. (1999) Hierarchical gene regulatory systems arising from fortuitous gene associations: controlling quorum sensing by the opine regulon in *Agrobacterium*. *Mol Microbiol* **32**: 1077-1089.

Platt, T.G., and Fuqua, C. (2010) What's in a name? The semantics of quorum sensing. *Trends Microbiol* **18**: 383-387.

Qin, Y., Luo, Z.Q., Smyth, A.J., Gao, P., Beck von Bodman, S., and Farrand, S.K. (2000) Quorum-sensing signal binding results in dimerization of TraR and its release from membranes into the cytoplasm. *Embo J* **19**: 5212-5221.

Qin, Y., Luo, Z.Q., and Farrand, S.K. (2004) Domains formed within the N-terminal region of the quorum-sensing activator TraR are required for transcriptional

activation and direct interaction with RpoA from *Agrobacterium*. *J Biol Chem* **279**: 40844-40851.

Qin, Y., Keenan, C., and Farrand, S.K. (2009) N- and C-terminal regions of the quorum-sensing activator TraR cooperate in interactions with the alpha and sigma-70 components of RNA polymerase. *Mol Microbiol* **74**: 330-346.

Ramirez-Romero, M.A., Soberon, N., Perez-Oseguera, A., Tellez-Sosa, J., and Cevallos, M.A. (2000) Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. *J Bacteriol* **182**: 3117-3124.

Ramirez-Romero, M.A., Tellez-Sosa, J., Barrios, H., Perez-Oseguera, A., Rosas, V., and Cevallos, M.A. (2001) RepA negatively autoregulates the transcription of the *repABC* operon of the *Rhizobium etli* symbiotic plasmid basic replicon. *Mol Microbiol* **42**: 195-204.

Ravin, N.V., Kuprianov, V.V., Gilcrease, E.B., and Casjens, S.R. (2003) Bidirectional replication from an internal ori site of the linear N15 plasmid prophage. *Nucleic Acids Res* **31**: 6552-6560.

Reisenauer, A., and Shapiro, L. (2002) DNA methylation affects the cell cycle transcription of the *CtrA* global regulator in *Caulobacter*. *Embo J* **21**: 4969-4977.

Rezzonico, F., and Duffy, B. (2008) Lack of genomic evidence of AI-2 receptors suggests a non-quorum sensing role for *luxS* in most bacteria. *BMC Microbiol* **8**: 154.

Ringgaard, S., van Zon, J., Howard, M., and Gerdes, K. (2009) Movement and equipositioning of plasmids by ParA filament disassembly. *Proc Natl Acad Sci U S A* **106**: 19369-19374.

Rodionov, O., Lobocka, M., and Yarmolinsky, M. (1999) Silencing of genes flanking the P1 plasmid centromere. *Science* **283**: 546-549.

Schumacher, M.A. (2007) Structural biology of plasmid segregation proteins. *Curr Opin Struct Biol* **17**: 103-109.

Schumacher, M.A. (2008) Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. *Biochem J* **412**: 1-18.

Shaheen, S.M., Ouimet, M-C, Marczynski, G.T. (2009) Comparative analysis of *Caulobacter* chromosome replication origins. *Microbiology* **155**: 1215-1225.

Slater, S.C., Goldman, B.S., Goodner, B., Setubal, J.C., Farrand, S.K., Nester, E.W., Burr, T.J., Banta, L., Dickerman, A.W., Paulsen, I., Otten, L., Suen, G., Welch, R., Almeida, N.F., Arnold, F., Burton, O.T., Du, Z., Ewing, A., Godsy, E., Heisel, S., Houmiel, K.L., Jhaveri, J., Lu, J., Miller, N.M., Norton, S., Chen, Q., Phoolcharoen, W., Ohlin, V., Ondrusek, D., Pride, N., Stricklin, S.L., Sun, J., Wheeler, C., Wilson, L., Zhu, H., and Wood, D.W. (2009) Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol* **191**: 2501-2511.

Smillie, C., Garcillan-Barcia, M.P., Francia, M.V., Rocha, E.P., and de la Cruz, F. (2010) Mobility of plasmids. *Microbiol Mol Biol Rev* **74**: 434-452.

Smith, E.F., and Townsend, C.O. (1907) A Plant-Tumor of Bacterial Origin. *Science* **25**: 671-673.

Soberon, N., Venkova-Canova, T., Ramirez-Romero, M.A., Tellez-Sosa, J., and Cevallos, M.A. (2004) Incompatibility and the partitioning site of the *repABC* basic replicon of the symbiotic plasmid from *Rhizobium etli*. *Plasmid* **51**: 203-216.

Subramoni, S., and Venturi, V. (2009) LuxR-family 'solos': bachelor sensors/regulators of signalling molecules. *Microbiology* **155**: 1377-1385.

Sukupolvi, S., and O'Connor, C.D. (1990) TraT lipoprotein, a plasmid-specified mediator of interactions between gram-negative bacteria and their environment. *Microbiol Rev* **54**: 331-341.

Surtees, J.A., and Funnell, B.E. (2001) The DNA binding domains of P1 ParB and the architecture of the P1 plasmid partition complex. *J Biol Chem* **276**: 12385-12394.

Tabata, S., Hooykaas, P.J., and Oka, A. (1989) Sequence determination and characterization of the replicator region in the tumor-inducing plasmid pTiB6S3. *J Bacteriol* **171**: 1665-1672.

Tanaka, T., Ishida, H., and Maehara, T. (2005) Characterization of the replication region of plasmid pLS32 from the Natto strain of *Bacillus subtilis*. *J Bacteriol* **187**: 4315-4326.

Thomas, C.M., and Smith, C.A. (1987) Incompatibility group P plasmids: genetics, evolution, and use in genetic manipulation. *Annu Rev Microbiol* **41**: 77-101.

Thomas, C.M. (2000) Paradigms of plasmid organization. *Mol Microbiol* **37**: 485-491.

Tsai, C.S., and Winans, S.C. (2010) LuxR-type quorum-sensing regulators that are detached from common scents. *Mol Microbiol* **77**: 1072-1082.

Turner, S.L., Rigottier-Gois, L., Power, R.S., Amarger, N., and Young, J.P. (1996) Diversity of *repC* plasmid-replication sequences in *Rhizobium leguminosarum*. *Microbiology* **142 (Pt 7)**: 1705-1713.

Van Melderen, L., and Saavedra De Bast, M. (2009) Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genet* **5**: e1000437.

Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R., Neddermann, P., and Marco, S.D. (2002) The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *Embo J* **21**: 4393-4401.

Vecchiarelli, A.G., Han, Y.W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertumpfel, C., Funnell, B.E., and Mizuuchi, K. (2010) ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Mol Microbiol* **78**: 78-91.

Venkova-Canova, T., Soberon, N.E., Ramirez-Romero, M.A., and Cevallos, M.A. (2004) Two discrete elements are required for the replication of a *repABC* plasmid: an antisense RNA and a stem-loop structure. *Mol Microbiol* **54**: 1431-1444.

Wagner-Dobler, I., Ballhausen, B., Berger, M., Brinkhoff, T., Buchholz, I., Bunk, B., Cypionka, H., Daniel, R., Drepper, T., Gerdts, G., Hahnke, S., Han, C., Jahn, D., Kalhoefer, D., Kiss, H., Klenk, H.P., Kyrpides, N., Liebl, W., Liesegang, H., Meincke, L., Pati, A., Petersen, J., Piekarski, T., Pommerenke, C., Pradella, S., Pukall, R., Rabus, R., Stackebrandt, E., Thole, S., Thompson, L., Tielen, P., Tomasch, J., von Jan, M., Wanphrut, N., Wichels, A., Zech, H., and Simon, M. (2010) The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. *Isme J* **4**: 61-77.

Watanabe, E., Inamoto, S., Lee, M.H., Kim, S.U., Ogua, T., Mori, H., Hiraga, S., Yamasaki, M., and Nagai, K. (1989) Purification and characterization of the sopB gene product which is responsible for stable maintenance of mini-F plasmid. *Mol Gen Genet* **218**: 431-436.

Waters, C.M., and Bassler, B.L. (2005) Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* **21**: 319-346.

Weaver, K.E. (2007) Emerging plasmid-encoded antisense RNA regulated systems. *Curr Opin Microbiol* **10**: 110-116.

White, C.E., and Winans, S.C. (2005) Identification of amino acid residues of the *Agrobacterium tumefaciens* quorum-sensing regulator TraR that are critical for positive control of transcription. *Mol Microbiol* **55**: 1473-1486.

White, C.E., and Winans, S.C. (2007) Cell-cell communication in the plant pathogen *Agrobacterium tumefaciens*. *Philos Trans R Soc Lond B Biol Sci* **362**: 1135-1148.

Whitehead, N.A., Barnard, A.M., Slater, H., Simpson, N.J., and Salmond, G.P. (2001) Quorum-sensing in Gram-negative bacteria. *FEMS Microbiol Rev* **25**: 365-404.

Winans, S.C. (1992) Two-way chemical signaling in *Agrobacterium*-plant interactions. *Microbiol Rev* **56**: 12-31.

Wion, D., and Casadesus, J. (2006) N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat Rev Microbiol* **4**: 183-192.

Wood, D.W., Setubal, J.C., Kaul, R., Monks, D.E., Kitajima, J.P., Okura, V.K., Zhou, Y., Chen, L., Wood, G.E., Almeida, N.F., Jr., Woo, L., Chen, Y., Paulsen, I.T., Eisen, J.A., Karp, P.D., Bovee, D., Sr., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M.J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Romero, P., Gordon, D., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z.Y., Dolan, M., Chumley, F., Tingey, S.V., Tomb, J.F., Gordon, M.P., Olson, M.V., and Nester, E.W. (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317-2323.

Zaenen, I., Van Larebeke, N., Van Montagu, M., and Schell, J. (1974) Supercoiled circular DNA in crown-gall inducing *Agrobacterium* strains. *J Mol Biol* **86**: 109-127.

Zhang, R.G., Pappas, T., Brace, J.L., Miller, P.C., Oulmassov, T., Molyneaux, J.M., Anderson, J.C., Bashkin, J.K., Winans, S.C., and Joachimiak, A. (2002) Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. *Nature* **417**: 971-974.

Zhu, J., and Winans, S.C. (1999) Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells. *Proc Natl Acad Sci U S A* **96**: 4832-4837.

Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J., Farrand, S.K., and Winans, S.C. (2000) The bases of crown gall tumorigenesis. *J Bacteriol* **182**: 3885-3895.

Zhu, J., and Winans, S.C. (2001) The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. *Proc Natl Acad Sci U S A* **98**: 1507-1512.

CHAPTER TWO

Dimerization of the Quorum Sensing Transcription Factor TraR Enhances Resistance to Cytoplasmic Proteolysis[2]

2.1. Summary

TraR is a LuxR-type quorum sensing protein encoded by Ti plasmid of *Agrobacterium tumefaciens*. TraR requires the pheromone 3-oxooctanoylhomoserine lactone (OOHL) for biological activity, and is dimeric both in solution and when bound to DNA. Dimerization is mediated primarily by two alpha helices, one in the N-terminal OOHL binding domain, and the other in the C-terminal DNA binding domain. Each of these helices forms a parallel coiled coil with the identical helix of the opposite subunit. We have previously shown that OOHL is essential for resistance to proteolysis, and here we asked whether dimerization is also required for protease resistance. We constructed a series of site-directed mutations at the dimer interface, and tested these mutants for activity *in vivo*. Alteration of residues A149, A150, A153, A222 and I229 completely abolished activity, while alteration of three other residues also caused significant defects. All mutants were tested for dimerization as well as for specific DNA binding. The cellular abundance of these proteins was measured using western immunoblots and OOHL-sequestration, while the half-life was measured by pulse-chase radiolabelling. We found a correlation between defects in *in vivo* activity, *in vitro* dimerization, DNA binding, intracellular abundance, and protease-resistance. We conclude that dimerization of TraR enhances resistance to cellular proteases.

---

[2] Chapter 2 is modified from a published manuscript with the permission from the co-author and the publisher. Pinto, U.M. and Winans, S.C. (2009) Dimerization of the quorum sensing transcription factor TraR enhances resistance to cytoplasmic proteolysis. *Mol Micro* **73**: 32-42.

2.2. Introduction

The TraR protein of *Agrobacterium tumefaciens* is a member of the LuxR family of transcription factors, and is a receptor of N-acylhomoserine lactones (AHLs). AHLs act as pheromones in intercellular communication (Eberhard, 1972). AHLs are generally synthesized by a cognate LuxI-type protein (Pappas et al., 2004). These pheromones accumulate in a population density-dependent manner and form complexes with their receptors when they reach a threshold concentration. LuxR/LuxI-type regulatory systems therefore allow bacteria to monitor their cell population densities, and to coordinate their behavior in a process referred to as quorum sensing. This phenomenon was first discovered in the marine bacterium *Vibrio fischeri*, whose LuxR and LuxI proteins regulate genes required for bioluminescence. Homologous systems have also been identified in many proteobacteria where they regulate diverse activities, including pathogenesis, biofilm formation, horizontal DNA transfer and the production of secondary metabolites (Waters and Bassler, 2005).

TraR regulates genes that are required for vegetative replication (*rep* genes) and conjugal transfer (*tra* and *trb* genes) of the tumour-inducing (Ti) plasmid (Fuqua and Winans, 1994; Li and Farrand, 2000; Pappas and Winans, 2003). TraR activity requires N-3-oxooctanoyl-l-homoserine lactone (OOHL), which is synthesized by the Ti plasmid-encoded TraI protein. TraR subunits form dimers that bind to dyad-symmetrical DNA sequences. It has two domains, an N-terminal OOHL-binding domain and a C-terminal DNA binding domain (Pappas et al., 2004). OOHL is required for TraR to become resistant to proteolysis *in vivo*, and is thought to act as a scaffold for TraR folding (Zhu and Winans, 1999, 2001).

The structure of TraR bound to OOHL and *tra* box DNA has been solved by X-ray crystallography (Vannini et al., 2002; Zhang et al., 2002). The N-terminal and

C-terminal domains both contribute to protein dimerization. The N-terminal domain dimerizes chiefly through α-helix 9 of each subunit (the longest α-helix in the protein), forming a coiled-coil (Figure 1A). The C-terminal domain dimerizes through α-helix 13 of each subunit, which form a second coiled-coil (Fig. 1B). Additional contacts include ionic interactions between the C-terminal carboxylate of each subunit and Arg230 of the opposite subunit.

Previous studies have shown that OOHL-mediated folding is essential for resistance to cytoplasmic proteases (Zhu and Winans, 1999, 2001). However, it was not known whether fully folded TraR monomers were protease resistant, or whether dimerization was also required. Cellular proteases are generally thought to detect denatured proteins by virtue of hydrophobic residues that would lie within the hydrophobic core of a properly folded protein (Wickner et al., 1999). The interface between TraR monomers is quite hydrophobic, and dimers are stabilized in large part by hydrophobic interactions. TraR monomers would therefore probably present a large hydrophobic surface, which might be recognized by proteases, and lead to proteolysis. Dimerization would sequester these hydrophobic patches, and could thereby block proteolysis. It therefore seemed plausible that protection from proteolysis might require not only binding of OOHL, but also dimerization of TraR monomers. If so, then mutations that block dimerization would be predicted to be more susceptible to proteolysis. In this study, TraR alleles were isolated that block biological activity, protein dimerization and DNA binding *in vitro*. These mutants showed an increased rate of turnover compared to wild type TraR.

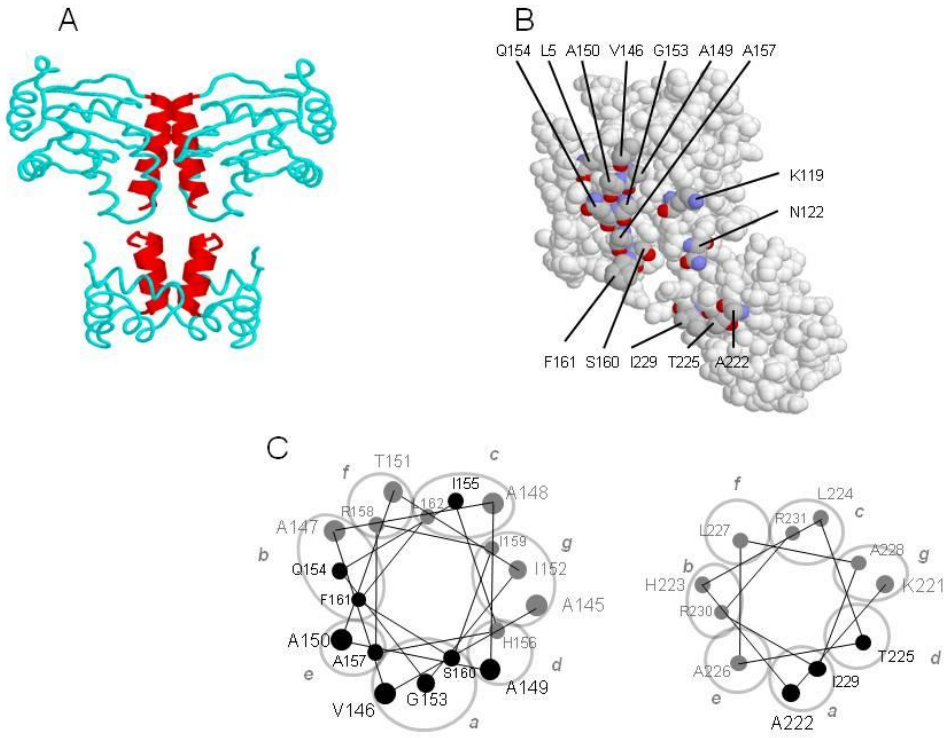


Figure 2.1. Dimerization interface of TraR.
A. Two alpha helices in the N-terminal domain form a coiled coil, and two additional alpha helices in the C-terminal domain form a second coiled coil. These two coiled coils provide the majority of the dimerization interface.
B. A monomer of TraR with the amino acid residues that were mutated in this study.
C. Helical wheel projection of alpha-helix 9 and alpha-helix 13 of TraR respectively. Residues depicted in black were altered by site-directed mutagenesis in this study. Positions *a* to *g* on the helical wheel projection are indicated.

2.3. Results

The goals of this study were to test amino acid residues that lie at the dimerization interface for their role in dimer formation and to determine whether defects in dimerization affect the rate at which these proteins are eliminated from the cells by proteolysis. Residues facing the dimerization interface and potentially mediating hydrophobic interactions were chosen for site-directed mutagenesis (Fig. 2.1). In addition, we altered residues K119 and N122 because they interact with residues D6 and D10 of the opposite monomer (Zhang et al., 2002) and could potentially aid in dimerization. Residues I155 and S160 were chosen based on a previous study that showed their contribution to dimerization (Luo et al., 2003). Three other mutations were included as controls that were unlikely to have a role in dimerization. Of these, alteration of residue R206 confers a DNA binding defective phenotype (White and Winans, 2007), while alteration of residue G123 confers a positive control phenotype (Luo and Farrand, 1999), and alteration of residue L5 is phenotypically silent (this work). Wherever possible, mutations were made that preserve interactions with neighboring amino acid residues.

Transcriptional activation

Each TraR mutant was tested *in vivo* for its ability to activate a PtraI-lacZ fusion using a broad range of OOHL concentrations (Table 2.1). Mutations A149E, A149V, A150E, A150V, G153E, Q154E, A222D, and I229Y caused strong defects in TraR activation, with transcriptional levels less than 15% of the wild type for all OOHL concentrations tested. Mutation of other residues had weaker effects on the transcription activity, especially at higher concentrations of OOHL (Table 2.1). Residues A149 and A150, which are located on the alpha-helix 9, were especially

critical for activity, as even conservative mutations strongly impaired transcriptional activity.

Table 2.1. *In vivo* activity at P*traI-lacZ* fusion of all point mutants relative to wild type TraR at different concentrations of OOHL.

OOHL concentration	0.1 nM	1nM	10nM	100nM
Wild type TraR	100.0	100.0	100.0	100.0
Vector Control	0.1	0.1	0.1	0.1
L5M	NT	100.0	100.0	100.0
K119A	28.6	56.7	40.2	61.1
N122A	0.1	2.9	21.9	29.3
G123R	NT	0.1	NT	0.1
V146E	5.9	56.4	60.5	75.7
V146A	45.2	41.3	39.1	33.9
A149E	0.2	0.1	0.1	0.1
A149V	0.1	0.3	1.1	1.6
A150E	0.1	1.3	6.2	8.0
A150V	0.2	0.1	0.5	1.5
G153E	0.1	0.1	0.1	0.1
G153A	17.1	84.8	105.1	90.0
Q154E	0.2	2.1	6.9	14.3
Q154A	0.6	7.9	56.0	40.3
I155F	NT	13.5	60.7	94.4
A157E	10.6	30.2	56.7	68.7
A157V	9.2	44.3	55.7	74.6
S160A	NT	92.0	99.0	89.0
S160F	NT	11.5	62.4	99.3
F161K	0.8	8.7	33.5	52.4
F161A	5.5	17.8	42.8	37.7
R206Q	NT	0.1	NT	NT
A222S	29.9	59.9	52.5	83.5
A222D	0.2	0.3	0.8	0.8
T225S	1.0	10.1	17.5	22.8
I229A	NT	35.0	NT	NT
I229Y	0.3	0.1	0.1	0.1

Activity was measured at P*traI-lacZ* fusion in NTL4 (pCEW260) with vector control (pPZP201), wild-type *traR* (pYC335) or each point mutant. Data represent the average of at least 3 repetitions compared to wild type TraR values that were set to 100%. Wild type TraR expression values were about 700, 2100, 2600 and 3000 units of β-galactosidase activity at 0.1 nM, 1 nM, 10 nM and 100 nM of OOHL, respectively. NT – Not tested.

Assays for TraR dimerization *in vitro*

The assays described above measured the ability of TraR mutants to activate transcription *in vivo*, but did not directly address the role of the mutated residues in dimerization. We therefore developed assays that directly evaluate the ability of these mutant proteins to form dimers. These assays involved the construction of fusions between these mutants and the maltose binding protein (MBP) of *E. coli*. It was previously shown that native TraR can form heterodimers with MBP-TraR fusions (Zhu and Winans, 2001). MBP-TraR fusion proteins containing point mutations in TraR were combined with wild type TraR. After allowing time for heterodimer formation, MBP-TraR fusion proteins were purified using an amylose-agarose affinity resin. The bound fractions were eluted using maltose and examined by SDS-PAGE gels for the presence of native TraR. In these assays, only the fusion protein contains the point mutation of interest, while the native TraR has a wild type sequence. In the *in vivo* assays described above, both subunits of the dimer contained identical mutations.

As expected, a fusion protein containing wild-type TraR sequences efficiently bound native TraR protein (Fig. 2.2). In contrast, most of the fusion proteins containing TraR mutations were defective in binding native TraR (Fig. 2.2, Table 2.2). In many cases, a correlation was found between defects in *in vivo* activity and heterodimer formation (Table 2.2). Residues A149, A150, G153, Q154, and I229 were especially strongly defective in both assays. A few mutant proteins did not follow this pattern, in that they were strongly defective in activity, yet had subtle defects in dimerization (Table 2.1 and 2.2, mutants G123R, R206Q, and T225S).

These mutants were judged to be defective in some other TraR property such as positive control and DNA binding.

Table 2.2. Correlation between accumulation, DNA binding, dimerization, and OOHL retention.

	Accumulation[1]	DNA binding[2]	Heterodimer Formation [1]	OOHL Retention [1]
Wild type TraR	(100)	Y	(100)	100
Vector Control	(0)	N	(0)	(0)
L5M	100	Y	99	100
K119A	82	Y	94	89
N122A	28	Y	28	35
G123R	51	Y	49	67
V146A	110	Y	61	104
V146E	65	Y	52	87
A149E	44	N	12	38
A149V	45	N	19	33
A150E	60	N	14	55
A150V	47	N	15	52
G153A	63	Y	61	66
G153E	33	N	19	39
Q154A	45	N	27	45
Q154E	36	N	25	28
I155F	43	Y	50	NT
A157E	48	Y	48	65
A157V	43	Y	50	75
S160A	100	Y	100	NT
S160F	55	Y	69	NT
F161A	82	Y	49	73
F161K	44	Y	39	57
R206Q	105	N	71	NT
A222D	38	N	29	32
A222S	66	Y	53	60
T225S	39	N	43	35
I229A	NT	Y	NT	33
I229Y	41	N	23	28

[1] Values represent the average of at least 3 repetitions. Wild type levels were set to 100%. NT – not tested. Accumulation was measured by semi-quantitative western blots using strain NTL4 (pCEW260)(pYC335) or its derivatives carrying each *traR* mutant, dimerization was determined by the ability of a MBP-TraR fusion carrying each point mutation to form heterodimers with native TraR, and OOHL retention assays was tested as previously described (Chai and Winans, 2004).

[2] DNA binding affinity of each TraR allele was evaluated by gel retardation assays using cleared cell lysates and *tra* box DNA repeated at least twice for each mutant. Y – TraR shifted DNA, N – TraR did not shift the DNA.

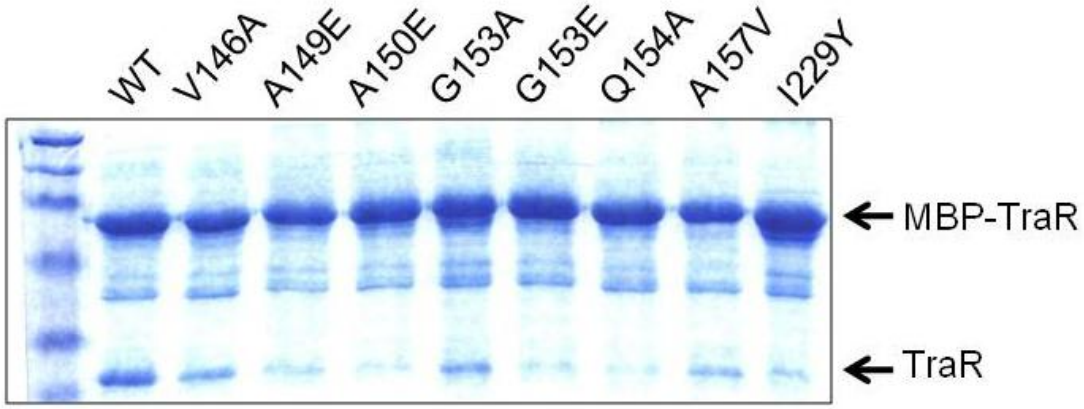


Figure 2.2. Dimerization of wild type TraR with MBP-TraR fusions containing point mutations at the subunit interface. Cell supernantants containing native TraR and various fusion proteins were combined and allowed to form heterodimers, then purified by amylose affinity chromatography. MBP-TraR fusions having wild type TraR sequence bind native TraR, while some MBP-TraR fusions having mutant TraR sequences fail to bind TraR.

Assays for binding to DNA fragments containing TraR binding sites

It is well established that TraR heterodimers containing just one DNA binding domain fail to bind to *tra* box DNA sequences (Chai et al., 2001; Oger et al., 1998; Zhu and Winans, 1998). This indicates that TraR monomers do not bind DNA with high affinity and that dimerization is essential for high affinity DNA binding. To test whether defects in dimerization lead to defects in DNA binding, we conducted electrophoretic mobility shift assays (EMSA) using wild type or mutant proteins and a DNA fragment that contains a consensus *tra* box. As expected, virtually all mutants that had strong defects in transcription activation and dimerization were defective in DNA binding (Fig. 2.3, Table 2.2), while mutants with subtle phenotypes in the former assays were proficient in DNA binding. An exception to this pattern was found in mutant N122A, which was defective in *in vivo* activity, yet showed wild-type affinity for *tra* box DNA. We conclude that this mutant is defective in positive

control. We have identified other residues that lie close to N122 that are also required for positive control (Costa et al, 2009).

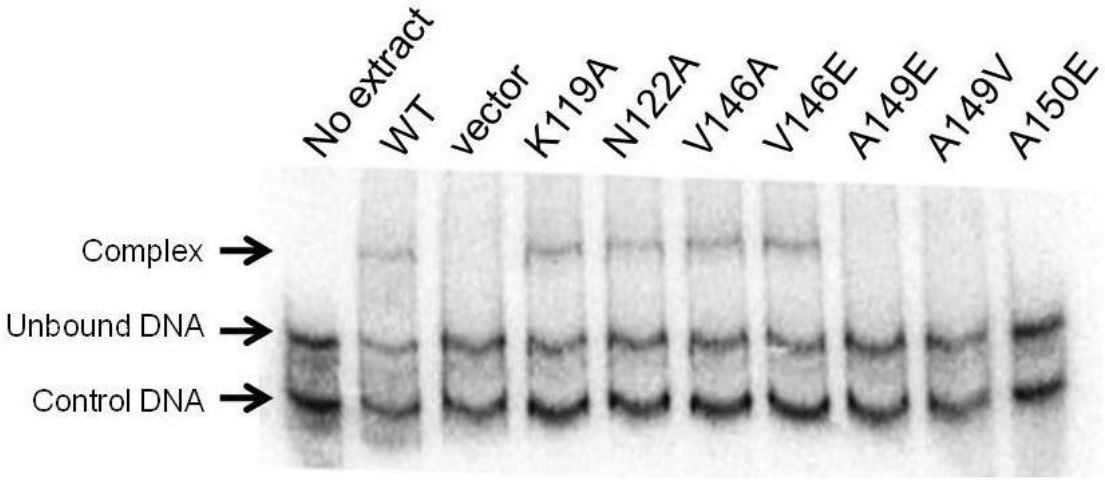


Figure 2.3. Electrophoretic mobility shift assays using clarified cell extracts containing wild type or mutant TraR proteins and a DNA fragment containing a consensus TraR binding site. Control DNA lacks a binding site for TraR. The amount of cell extracts added was normalized for TraR abundance using Western immunoblots.

In vivo accumulation and stability of dimerization mutants

The central question addressed in this study is whether dimerization of TraR is required for resistance to proteolysis. We addressed this question in three ways, two of which involve measurements of TraR abundance, and one of which directly measures the half-life of these mutant proteins. TraR abundance was assayed, first, by measuring the ability of whole cells expressing wild type or mutant protein to sequester exogenously provided OOHL. TraR mutants that are rapidly degraded or that cannot fold are defective in this assay (Chai and Winans, 2004). All dimerization mutants were proficient in binding OOHL, though the levels of OOHL binding were lower when compared to the wild type TraR (Table 2.2).

We also assayed for the intracellular abundance of each mutant protein by semi-quantitative western immunoblots. All mutant TraR proteins were readily detected immunologically, even those having very strong defects in transcription activation, dimerization, and DNA binding (Fig. 2.4, Table 2.2). There was little if any correlation between accumulation in this assay and the ability to form dimers, as measured above. These data would support the idea that dimerization does not play a critical role in resistance to proteolysis, though it might not detect subtle effects in accumulation.

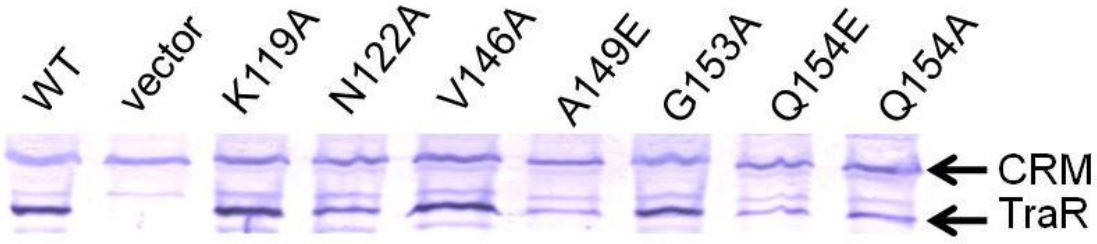


Figure 2.4. Western immunoblots showing the abundance of wild type and mutant TraR proteins in clarified cell extracts. Cells were cultured in the presence of 10 nM OOHL. CRM is cross-reacting material, which was used to ensure that equivalent amounts of total protein was added to each lane. Strain carrying pPZP201 vector was used as a negative control while pYC335 served as a positive control for TraR.

We used pulse-chase experiments to evaluate the rate at which each mutant is degraded *in vivo*, as previously done using wild type protein (Zhu and Winans, 2001). As a negative control, a strain growing in the absence of OOHL showed very little accumulation of soluble TraR (Fig. 2.5), and the detectable TraR was rapidly degraded. The same strain cultivated in the presence of OOHL showed little or no TraR degradation. Mutant TraR proteins that have strong defects in dimerization, DNA binding, and activity were degraded more rapidly than the wild type protein in the presence of OOHL, but far more slowly than the wild type in the absence of

OOHL. We conclude that blocks in dimerization lead to decreases in protease resistance, but that this effect is more subtle than the requirement of OOHL.

2.4. Discussion

The crystal structure of TraR-OOHL-DNA ternary complexes has enabled a series of studies about the roles of particular amino acid residues in binding of OOHL, in decoding of *tra* box DNA, in binding to RNA polymerase, and now in the dimerization of the TraR subunits (Chai and Winans, 2004; White and Winans, 2005, 2007). Each of these studies provides an ever-sharper picture of how TraR subunits interact with their molecular environment. The results obtained in this work corroborate a model (Fig. 2.6) in which OOHL binds to and stabilizes TraR, allowing time for monomers to form dimers before they get degraded. OOHL binding and dimerization greatly enhances resistance to degradation.

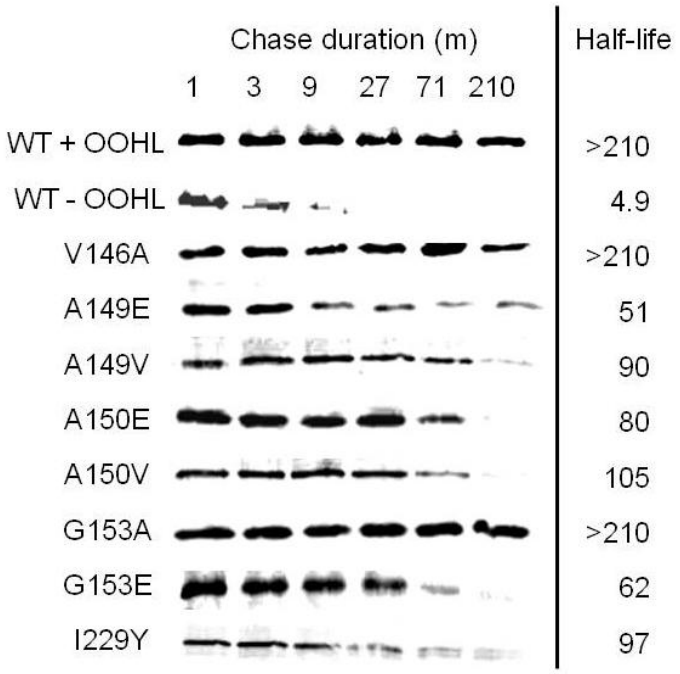


Figure 2.5. Pulse chase of TraR point mutants in *E. coli*. Cells expressing wild type or TraR point mutants from a phage T7 promoter were treated with rifampicin to block host transcription, then treated with [^{35}S]methionine for 1 min, followed by addition of excess nonlabeled methionine. OOHL was provided prior to addition of radiolabel. At the indicated intervals after the addition of nonlabeled methionine, aliquots were frozen at -80°C to terminate proteolysis, then thawed and size-fractionated by SDS-PAGE. Radiolabel was quantitated using a Storm PhosphorImager. Calculated TraR half-lives are indicated at the right of each panel and represent the average of two independent experiments.

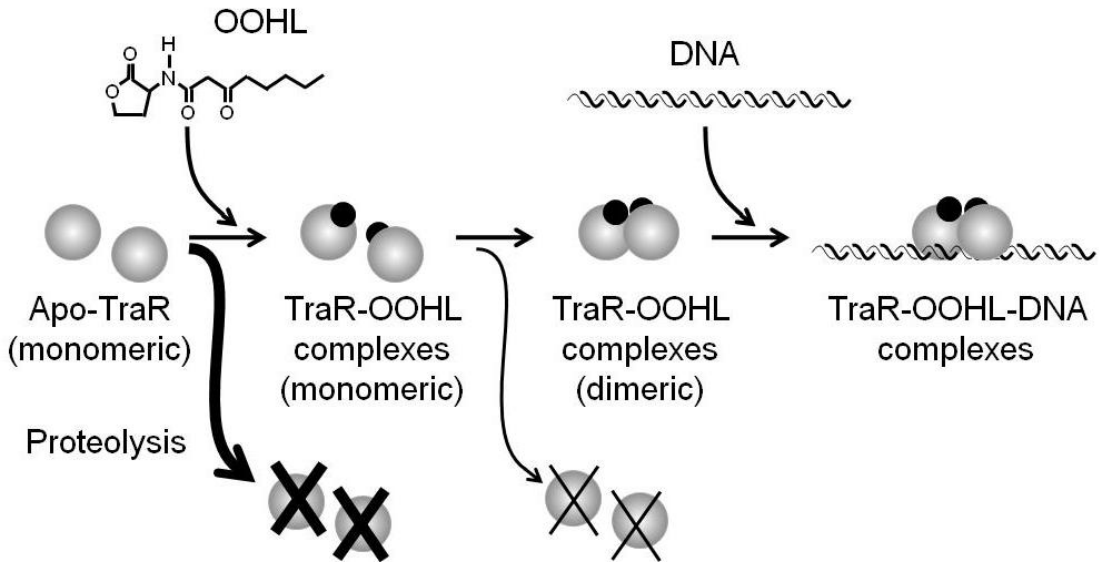


Figure 2.6. TraR folding and dimerization model. TraR monomers that fail to bind OOHL after synthesis are immediately targeted for proteolysis (Zhu and Winans, 1999, 2001), while TraR-OOHL complexes that fail to dimerize are also targeted for proteolysis, though not as rapidly as apo-TraR. TraR-OOHL dimers are competent to bind *tra* box DNA with high affinity and specificity.

By altering all residues at the subunit interface, we were able to evaluate the contribution of each to dimer formation and overall function. Residues A149, A150, and G153, all from helix 9, were critical for dimerization, as the mutations A149V, A149E, A150V, A150E, and G153E caused very strong defects in dimer formation (Fig. 2.2, Table 2.2). This could be due to the larger size of the mutant residues, and to the lack of available unoccupied space in this region of the dimer interface. Elsewhere in helix 9, mutations V146A, V146E, A157V and A157E caused relatively mild defects, even though the wild type residues make direct contacts with the opposite subunit. The structure suggests sufficient space to accommodate bulkier residues. Helix 13 must also play a role in dimerization, as mutations A222D and I229Y caused severe defects in activity, DNA binding, and heterodimer formation

(Tables 2.1 and 2.2), though more conservative mutations had milder defects. Overall the results indicate that positions a and d on the helical wheel projection (which corresponds to residues 149, 153, 222, 225, and 229 on TraR, Fig. 2.1) show specific defects in protein dimerization (Lupas, 1996). A study with the *E. coli* FNR transcription factor also showed that residues located at similar positions on the helical wheel projection were the main ones directly involved in FNR dimerization (Moore and Kiley, 2001).

Though this study focused on the use of TraR mutants, we hope to make inferences about the wild type protein. Does wild type TraR ever exist as a monomer, and if so, what are its properties? At the earliest stages of quorum sensing induction, some cells could have as few as one TraR monomer. Even in cells with more than one monomer, newly synthesized TraR monomers must require a short time interval to dimerize, and even then, dimers exist in a dynamic equilibrium with monomers (Chai et al., 2001; Oger et al., 1998; Zhu and Winans, 1998). Would such monomers be protease-sensitive? For several reasons, we believe that the mutants we have described are a fair approximation of wild type monomers. First, wherever possible, mutations were chosen that preserve interactions with neighboring amino acid residues and change only the surface of the protein. This was done to try to minimize defects in tertiary structure. Second, most of the mutations described here did not perturb the sequestration of OOHL *in vivo*, suggesting that their overall tertiary structure was preserved. Third, there was a correlation between the severity in the dimerization defect and protein instability. Those mutations with the strongest defects in dimerization (Fig. 2.2) tended to have the strongest defects in stability (Fig. 2.5). Conversely, weaker defects in dimerization tended to be correlated with weaker defects in stability. These data strongly suggest that blocks in dimerization lead

directly to decreased half-life, and that these mutants are informative about the properties of wild type monomeric protein.

TraR monomers would likely have a considerable hydrophobic surface exposed to solvent, which could in principle target the protein for proteolysis. It is conceivable however, that the hydrophobic patches of the two domains could interact with each other, possibly forming an antiparallel coiled coil. If so, this might shield both domains from proteases, and possibly increase the solubility of the protein. If that were true, the formation of a TraR dimer would require the dissociation of the N-terminal and C-terminal domains of each monomer before they could interact with their counterparts of the opposite subunit.

Dimerization of transcription factors is often mediated by parallel or antiparallel helices (Lupas, 1996). For instance, the catabolite activator protein (CAP) of *E. coli* dimerizes via two parallel α-helices that extend from residues 108 to 137 of each subunit (Joung et al., 1995; Schultz et al., 1991). The *E. coli* protein FNR (a CAP homolog), has been mutagenized in the corresponding dimerization interface, and the resulting mutants invariably showed defects in dimerization and in DNA binding (Moore and Kiley, 2001). FNR dimerization occurs only in anaerobic conditions, in which a (4Fe-4S)2+ cluster in each subunit is reduced to a (2Fe-2S)2+ form (Kiley and Beinert, 1998). The former is thought to distort the dimerization helix, thereby blocking dimer formation (Kiley and Beinert, 1998). The receiver domain of the *E. coli* PhoB protein has been resolved in the active and inactive conformations (Bachhawat et al., 2005). Dimerization of the inactive form is mediated primarily by α helix 1 of each subunit, while dimerization of the active form is mediated by α-helix 4, α-strand 5, and α-helix 5 of each subunit. In the former state, the two DNA binding domains of the dimer cannot associate, while in the latter state, they are brought into close proximity, allowing them to bind DNA. The

transcription factor BmrR of *Bacillus subtilis* consists of an N-terminal DNA binding domain and a C-terminal ligand binding domain that are connected by a long α helix (residues 77-119). This helix forms a long antiparallel coiled coil with its counterpart from the opposite subunit (Heldwein and Brennan, 2001). This structure is thought to be conserved among other members of the MerR family. Two dimers of *E. coli* Lac repressor associate via four C-terminal α-helices, each 20 amino acids in length, with two antiparallel helices contributed from each dimer (Lewis et al., 1996). Each of these helices is connected to the rest of the subunit by a flexible linker, allowing considerable conformational freedom between the two dimers. The LysR-type regulator CbnR of *Ralstonia eutropha* consists of a N-terminal DNA binding domain fused to a C-terminal ligand binding domain via a long α helix that forms a parallel coiled coil with the opposite subunit (Muraoka et al., 2003). In contrast, some transcription factors, including Lrp of *E. coli*, PrgX of *Enterococcus facaelis*, dimerize by other sorts of secondary structures (de los Rios and Perona, 2007; Shi et al., 2005). We have not seen other reports showing that multimerization of any of these proteins plays a role in their stability.

2.5. Materials and methods

Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 2.3. *E. coli* strains were cultured in Luria broth (LB) or solid medium at 37°C (Miller, 1972). *A. tumefaciens* strains were cultured in AT minimal medium at 28°C (Cangelosi et al., 1991). Synthetic OOHL was provided by Dr. Anatol Eberhard (Cornell University). Antibiotics were added at the following concentrations: 100 μg ml^{-1} spectinomycin; 100 μg ml^{-1} kanamycin, and 100 μg ml^{-1} ampicillin. IPTG was added at 500 μM.

DNA manipulations

Recombinant DNA techniques were performed using established procedures (Sambrook and Russel, 2001). Plasmid DNA was isolated using QIAprep spin miniprep kits (Qiagen). DNA fragments generated by PCR or restriction digestion were gel purified using QIAquick Gel Extraction Kit (Qiagen). Restriction endonucleases were obtained from New England Biolabs and used according to methods described by the manufacturers. Plasmid DNA was introduced into *E. coli* and *A. tumefaciens* by electroporation (Cangelosi et al., 1991).

Site-directed mutagenesis of *traR* and cloning into different vectors

Site-directed mutagenesis of *traR* was performed by using a synthetic overlap extension PCR with a four-primer approach (Sambrook and Russel, 2001). For mutations on the N-terminal domain of TraR, a 978 bp fragment of pYC335 (Chai and Winans, 2004) was amplified using *Pfx* DNA Polymerase (Invitrogen) to include an unique EcoRI site (located upstream of *traR*) and a native SacII site (near codon 168 of *traR*). For mutations on the C-terminal domain of TraR, a 303 bp fragment of pYC335 was amplified as above to include a native SacII site (near codon 168 of *traR*) and a unique MfeI site (just downstream of the stop codon of *traR*). All oligonucleotides used in this study are listed in Table 2.4 and were obtained from Integrated DNA Technologies (Coralville, IA). The flanking primers were used in separate reactions with two different mutagenic primers that overlap at the mutation, using pYC335 as the template. These two PCR products were then combined and used as the template in a second round of PCR with the same flanking primers to generate the complete fragment. The second set of PCR products was digested with EcoRI and SacII (for N-terminal domain mutations) or SacII and MfeI (for C-terminal domain

mutations), and ligated to pYC335 digested with the same enzymes. Mutant sequences were confirmed by automated DNA sequencing.

The traR point mutants were also cloned into vectors pAFM02 and pJZ358. For the former plasmid, mutants that lie between codons 85 and 168 were cloned by digesting pYC335 derivatives with BstB1 and SacII enzymes. The digestion of pYC335 derivatives with the above enzymes generated four fragments of distinct sizes. The one of 306 bp contained the mutated codons that were ligated in pAFM02 digested with the same enzymes. For other mutants located downstream of the codon 168, the unique sites, in both plasmids, SacII and HindIII were used. For cloning into pJZ358, digestion was performed with HindIII and BbsI enzymes. A fragment of 867 bp originating from pYC335 derivatives starting at codon 24 (BbsI site) of traR and going 227 bp downstream of its stop codon (HindIII site) was used to replace a fragment of 922 bp from pJZ358 digested with the same enzymes.

Measurement of TraR activity *in vivo*

Bioassays of TraR activity were performed with each TraR mutant using NTL4(pCEW260)(pYC335) or derivatives of pYC335 carrying each of the traR mutants. Strains were cultured in AT minimal medium at 28°C to an OD_{600} of 0.3–0.4. Each culture was then diluted 20-fold into AT medium containing concentrations of 0.1, 1, 10 or 100 nM OOHL, incubated with vigorous aeration at 28°C for 8 h, and assayed for β-galactosidase activity (Miller, 1972). Experiments were performed in triplicate (with three different isolates for each strain).

Table 2.3. Strains and plasmids.

Strains	Relevant features	References
DH5α	*E. coli,* α-complementation	Stratagene
NTL4	*A. tumefaciens* C58, pTi less	(Luo *et al.*, 2001)
BL21/DE3	*E. coli B* Plac-gene 1 of bacteriophage T7	(Studier *et al.*, 1990)
KY2347	*E. coli* MG1655, Δ(*clpPX-lon*) 1196::*cat*	(Herman *et al.*, 1998)
Plasmids		
pPZP201	Broad-host-range cloning vector, Spc^R	Hajdukiewicz et al. (1994)
pMCSG9	PT7-his6-MBP-TEV, Amp^R	(Donnelly et al. (2006)
pJZ358	PT7-*traR*, Amp^R	Zhu and Winans (1999)
pYC335	P*lac-traR* in pPZP201	(Chai and Winans, 2004
pCEW250	pBluescriptSK+ with consensus *tra* box at P*traI*, Km^R	(White and Winans, 2007)
pCEW260	Consensus P*traI–lacZ* fusion, Km^R	(White and Winans, 2007)
pAFM02	*traR* cloned into pMCSG9	This study
pUP-5M to 229Y	*traR* point mutants cloned into pYC335	This study
pUP2-5M to 229Y	*traR* point mutants cloned into pAFM02	This study
pUP3-5M to 229Y	*traR* point mutants cloned into pJZ358	This study

Table 2.4. Oligonucleotide primers used in this study.

Oligonucleotide Name	DNA Sequence
Flanking primers – N-terminal region	
PT1	5’ – CTCACTCATTAGGCACCCCAG – 3’
TraRb	5’ – GTACAACGTGTAGGGCAACGC- 3’
Flanking primers – C-terminal region	
TraRa	5’ – CATTCCTTCGCACCACCCC – 3’
TraRb	5’ – GTACAACGTGTAGGGCAACGC – 3’
Mutagenic primers	
L5M-F	5’ – CAGCACTGGATGGACAAGCTG – 3’
L5M-R	5’- CAGCTTGTCCATCCAGTGCTG – 3’
K119A-F	5’ – AATACCCATCGCGACCGCCAACG – 3’
K119A-R	5’ – CGTTGGCGGTCGCGATGGGTATT – 3’
N122A-F	5’ – CAAGACCGCCGCCGGCTTTATGT – 3’
N122A-R	5’ – ACATAAAGCCGGCGGCGGTCTTG – 3’
V146A-F	5’ – GATCGATGCAGCCGCAGCCGCTG – 3’
V146A-R	5’ – CAGCGGCTGCGGCTGCATCGATC – 3’
V146E-F	5’ – GATCGATGCAGAAGCAGCCGCTG – 3’
V146E-R	5’ – CAGCGGCTGCTTCTGCATCGATC – 3’
A149E-F	5’ – AGTCGCAGCCGAGGCAACCATCG – 3’
A149E-R	5’ – CGATGGTTGCCTCGGCTGCGACT – 3’
A149V-F	5’ – AGTCGCAGCCGTTGCAACCATCG – 3’
A149V-R	5’ – CGATGGTTGCAACGGCTGCGACT – 3’
A150E-F	5’ – CGCAGCCGCTGAAACCATCGGGC – 3’
A150E-R	5’ – GCCCGATGGTTTCAGCGGCTGCG – 3’
A150V-F	5’ – CGCAGCCGCTGTAACCATCGGGC – 3’
A150V-R	5’ – GCCCGATGGTTACAGCGGCTGCG – 3’
G153A-F	5’ – TGCAACCATCGCGCAGATCCATG – 3’
G153A-R	5’ – CATGGATCTGCGCGATGGTTGCA – 3’
G153E-F	5’ – TGCAACCATCGAGCAGATCCATG – 3’
G153E-R	5’ – CATGGATCTGCTCGATGGTTGCA – 3’

Table 2.4 (continued)

Q154A-F	5’ – AACCATCGGGGCGATCCATGCCC – 3’
Q154A-R	5’ – GGGCATGGATCGCCCCGATGGTT – 3’
Q154E-F	5’ – AACCATCGGGGAGATCCATGCCC – 3’
Q154E-R	5’ – GGGCATGGATCTCCCCGATGGTT – 3’
I155F-F	5’ – ATCGGGCAGTTCCATGCCCGC – 3’
I155F-R	5’ – GCGGGCATGGAACTGCCCGAT – 3’
A157E-F	5’ – GCAGATCCATGAGCGCATCTCAT – 3’
A157E-R	5’ – ATGAGATGCGCTCATGGATCTGC – 3’
A157V-F	5’ – GCAGATCCATGTCCGCATCTCAT – 3’
A157V-R	5’ – ATGAGATGCGGACATGGATCTGC – 3’
S160A-F	5’ – CCGCATCGCATTCCTTCGCA – 3’
S160A-R	5’ – TGCGAAGGAATGCGATGCGG – 3’
S160F-F	5’ – GCCCGCATCTTCTTCCTTCGC – 3’
S160F-R	5’ – GCGAAGGAAGAAGATGCGGGC – 3’
F161A-F	5’ – CCGCATCTCAGCCCTTCGCACCA – 3’
F161A-R	5’ – TGGTGCGAAGGGCTGAGATGCGG – 3’
F161K-F	5’ – CCGCATCTCAAAACTTCGCACCA – 3’
F161K-R	5’ – TGGTGCGAAGTTTTGAGATGCGG – 3’
A222D-F	5’ – GCAGCAAGGACCATCTTACC – 3’
A222D-R	5’ – GGTAAGATGGTCCTTGCTGC – 3’
A222S-F	5’ – GCAGCAAGTCCCATCTTACC – 3’
A222S-R	5’ – GGTAAGATGGGACTTGCTGC – 3’
T225S-F	5’ – CCCATCTTTCCGCGCTCGCC – 3’
T225S-R	5’ – GGCGAGCGCGGAAAGATGGG – 3’
I229A	(White and Winans, 2005)
I229Y-F	5’ – GCTCGCCTACCGGCGGAAAC – 3’
I229Y-R	5’ – GTTTCCGCCGGTAGGCGAGC – 3’
traR cloning into pMCSG9	
traR-fusion-F	5’ – TACTTCCAATCCAATATGCAGCACTGGCT – 3’
traR-fusion-R	5’ – TTATCCACTTCCAATTCTCAGATGAGTTTCCG – 3’

Table 2.4 (continued)

To check correct cloning into pMCSG9	
pMCSG9-F	5’ – ACGAGGAAGAGTTGGCGAAAGATC – 3’
pMCSG9-R	5’ – TTAGAGGCCCCAAGGGGTTATGCTA – 3’
Oligonucleotides used in the gel shift assays	
Ptra-box For	5’ – GAATTCTATGTGCAGATCTGCACATAGC – 3’
Ptra-box Rev	5’ – GGATCAATACGACGAGCTCGAGGATCCAGC – 3’
Pcontrol gel shift For	5’ – CCGCTACAGGGCGCGTCC – 3’
Pcontrol gel shift rev	5’ – CCAATTCGCCCTATAGTG – 3’

Heterodimer formation assay

To measure the ability of TraR to form heterodimers with MBP-TraR fusions, clear supernatants containing MBP-TraR fusions expressed from *E. coli* BL21/DE3(pAFM02) or pAFM02 derivatives carrying each TraR mutant were combined with equal volumes of clear supernatants containing native TraR protein expressed from BL21/DE3(pJZ358). Combined supernatants were incubated for 2 h at 28°C to allow subunit exchange, and were applied to an amylose affinity chromatography column (New England Biolabs). Proteins were step eluted with buffer containing 10 mM maltose. Purified proteins were run on 12% SDS-PAGE gels and band intensities were quantified using ImageJ software (http://rsbweb.nih.gov/ij/). Cells expressing only TraR or only MBP-TraR respectively were used as negative and positive controls.

Electrophoretic mobility shift assays

To perform gel shift assays, clarified cell extracts were prepared from *E. coli* strain KY2347 (a *clp*, *lon* mutant) carrying pYC335 or its derivatives with each of the *traR* point mutants. Strains were cultured in LB broth containing 100 μg mL^{-1} of

spectinomycin in the presence of 200 nM OOHL at 28°C. When cultures reached OD_{600} of 0.2, IPTG was added to a final concentration of 500 μM to induce TraR expression, and incubation was followed by an additional 6 h. TraR abundance in the soluble fraction was estimated using Western immunoblots as previously described (White and Winans, 2005). A 247 nucleotide PCR fragment containing a consensus *tra* box was obtained from plasmid pCEW250 using primers Ptra-box For and Ptra-box Rev (Table 2.4). A negative control fragment of 211 nucleotides in length was PCR amplified from plasmid pCEW250 with primers Pcontrol gel shift For and Pcontrol gel shift Rev. Both fragments were end-labelled with [γ-32P]-dATP (Pelkin Elmer) using T4 polynucleotide kinase (New England Biolabs). Binding reactions were performed as previously described (White and Winans, 2005), with equivalent amounts of soluble TraR in each reaction. Gel shift assays were performed with independent clarified lysates in duplicate for each strain, and data were analyzed using a Storm B840 Phosphorimager.

Measurements of TraR abundance in *A. tumefaciens*

The abundance of each TraR allele was determined using strain NTL4(pCEW160)(pYC335) or derivatives of pYC335 carrying each of the traR mutants. Strains were cultured in AT at 28°C with 10 nM OOHL, cells were harvested at mid-log phase, and westerns were performed as described previously (White and Winans, 2005). Westerns were prepared with independent cell lysates at least twice for each strain. Data were analyzed using IMAGEJ and normalized against cross-reacting material in each lane.

OOHL sequestration assay by TraR in whole cells.

A. tumefaciens strain NTL4(pCEW260)(pYC335) or derivatives of pYC335 carrying each TraR mutant allele were used for OOHL sequestration assays as previously described (Chai et al., 2001). All assays were performed twice with independent cultures.

TraR stability in *E. coli*

The measurement of TraR turnover for key TraR point mutants was determined using strain BL21/DE3(pJZ358) (Zhu and Winans, 1999) or derivatives of pJZ358 carrying *traR* mutants. The experiments were performed as described previously (Zhu and Winans, 2001) except that OOHL was added 10 min before the addition of radiolabel.

Acknowledgements

The authors gratefully acknowledge Dr Anatol Eberhard (Cornell University) for the synthesis and purification of OOHL used in this study, and Ana Lidia Flores-Mireles for construction of pAFM02 plasmid. We also thank all members of our laboratory for helpful discussions and critical review of the manuscript. Funding for this study was provided by the National Institute of General Medical Sciences (GM41892). U.M.P. acknowledges the financial support of the Brazilian government through a fellowship grant from the Coordenação de Aperfeicoamento de Pessoal de Nível Superior (Capes).

REFERENCES

Bachhawat, P., Swapna, G.V., Montelione, G.T., and Stock, A.M. (2005) Mechanism of activation for transcription factor PhoB suggested by different modes of dimerization in the inactive and active states. *Structure* **13**: 1353-1363.

Chai, Y., Zhu, J., and Winans, S.C. (2001) TrlR, a defective TraR-like protein of *Agrobacterium tumefaciens*, blocks TraR function *in vitro* by forming inactive TrlR:TraR dimers. *Mol Microbiol* **40**: 414-421.

Chai, Y., and Winans, S.C. (2004) Site-directed mutagenesis of a LuxR-type quorum-sensing transcription factor: alteration of autoinducer specificity. *Mol Microbiol* **51**: 765-776.

Costa, E.D., Cho, H., and Winans, S.C. (2009) Identification of amino acid residues of the pheromone-binding domain of the transcription factor TraR that are required for positive control. *Mol Microbiol* **73**: 341-351.

de los Rios, S., and Perona, J.J. (2007) Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *J Mol Biol* **366**: 1589-1602.

Donnelly, M.I., Zhou, M., Millard, C.S., Clancy, S., Stols, L., Eschenfeldt, W.H., Collart, F.R., and Joachimiak, A. (2006) An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expr Purif* **47**: 446-454.

Eberhard, A. (1972) Inhibition and activation of bacterial luciferase synthesis. *J Bacteriol* **109**: 1101-1105.

Fuqua, W.C., and Winans, S.C. (1994) A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J Bacteriol* **176**: 2796-2806.

Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol Biol* **25**: 989-994.

Heldwein, E.E., and Brennan, R.G. (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409**: 378-382.

Herman, C., Thevenet, D., Bouloc, P., Walker, G.C., and D'Ari, R. (1998) Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). *Genes Dev* **12**: 1348-1355.

Joung, J.K., Chung, E.H., King, G., Yu, C., Hirsh, A.S., and Hochschild, A. (1995) Genetic strategy for analyzing specificity of dimer formation: *Escherichia coli* cyclic AMP receptor protein mutant altered in its dimerization specificity. *Genes Dev* **9**: 2986-2996.

Kiley, P.J., and Beinert, H. (1998) Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. *FEMS Microbiol Rev* **22**: 341-352.

Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G., and Lu, P. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**: 1247-1254.

Li, P.L., and Farrand, S.K. (2000) The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. *J Bacteriol* **182**: 179-188.

Luo, Z.Q., and Farrand, S.K. (1999) Signal-dependent DNA binding and functional domains of the quorum-sensing activator TraR as identified by repressor activity. *Proc Natl Acad Sci U S A* **96**: 9009-9014.

Luo, Z.Q., Clemente, T.E., and Farrand, S.K. (2001) Construction of a derivative of *Agrobacterium tumefaciens* C58 that does not mutate to tetracycline resistance. *Mol Plant Microbe Interact* **14**: 98-103.

Luo, Z.Q., Smyth, A.J., Gao, P., Qin, Y., and Farrand, S.K. (2003) Mutational analysis of TraR. Correlating function with molecular structure of a quorum-sensing transcriptional activator. *J Biol Chem* **278**: 13173-13182.

Lupas, A. (1996) Coiled coils: new structures and new functions. *Trends Biochem Sci* **21**: 375-382.

Moore, L.J., and Kiley, P.J. (2001) Characterization of the dimerization domain in the FNR transcription factor. *J Biol Chem* **276**: 45744-45750.

Muraoka, S., Okumura, R., Ogawa, N., Nonaka, T., Miyashita, K., and Senda, T. (2003) Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. *J Mol Biol* **328**: 555-566.

Oger, P., Kim, K.S., Sackett, R.L., Piper, K.R., and Farrand, S.K. (1998) Octopine-type Ti plasmids code for a mannopine-inducible dominant-negative allele of traR, the quorum-sensing activator that regulates Ti plasmid conjugal transfer. *Mol Microbiol* **27**: 277-288.

Pappas, K.M., and Winans, S.C. (2003) A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol Microbiol* **48**: 1059-1073.

Pappas, K.M., Weingart, C.L., and Winans, S.C. (2004) Chemical communication in proteobacteria: biochemical and structural studies of signal synthases and receptors required for intercellular 85ignaling. *Mol Microbiol* **53**: 755-769.

Sambrook, J., and Russel, D.W. (2001) *Molecular Cloning – A laboratory manual.* New York: Cold Spring Harbor Laboratory Press.

Schultz, S.C., Shields, G.C., and Steitz, T.A. (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* **253**: 1001-1007.

Shi, K., Brown, C.K., Gu, Z.Y., Kozlowicz, B.K., Dunny, G.M., Ohlendorf, D.H., and Earhart, C.A. (2005) Structure of peptide sex pheromone receptor PrgX and PrgX/pheromone complexes and regulation of conjugation in *Enterococcus faecalis*. *Proc Natl Acad Sci U S A* **102**: 18596-18601.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**: 60-89.

Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R., Neddermann, P., and Marco, S.D. (2002) The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. *Embo J* **21**: 4393-4401.

Waters, C.M., and Bassler, B.L. (2005) Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* **21**: 319-346.

White, C.E., and Winans, S.C. (2005) Identification of amino acid residues of the *Agrobacterium tumefaciens* quorum-sensing regulator TraR that are critical for positive control of transcription. *Mol Microbiol* **55**: 1473-1486.

White, C.E., and Winans, S.C. (2007) The quorum-sensing transcription factor TraR decodes its DNA binding site by direct contacts with DNA bases and by detection of DNA flexibility. *Mol Microbiol* **64**: 245-256.

Wickner, S., Maurizi, M.R., and Gottesman, S. (1999) Posttranslational quality control: folding, refolding, and degrading proteins. *Science* **286**: 1888-1893.

Zhang, R.G., Pappas, T., Brace, J.L., Miller, P.C., Oulmassov, T., Molyneaux, J.M., Anderson, J.C., Bashkin, J.K., Winans, S.C., and Joachimiak, A. (2002) Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. *Nature* **417**: 971-974.

Zhu, J., and Winans, S.C. (1998) Activity of the quorum-sensing regulator TraR of *Agrobacterium tumefaciens* is inhibited by a truncated, dominant defective TraR-like protein. *Mol Microbiol* **27**: 289-297.

Zhu, J., and Winans, S.C. (1999) Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters *in vitro* and decreases TraR turnover rates in whole cells. *Proc Natl Acad Sci U S A* **96**: 4832-4837.

Zhu, J., and Winans, S.C. (2001) The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. *Proc Natl Acad Sci U S A* **98**: 1507-1512.

CHAPTER THREE

Ti plasmid entry exclusion: Quorum Sensing Converts *Agrobacterium tumefaciens* from Phenotypically Female to Male[3]

3.1. Summary

Conjugative plasmids generally encode proteins that block the conjugative entry of identical or similar plasmids into the host cell, a phenomenon known as entry exclusion. Here we demonstrate that two Ti plasmids of *Agrobacterium tumefaciens* encode robust entry exclusion functions. Two proteins, TrbJ and TrbK, can each mediate entry exclusion, and act synergistically. The *trbJ* and *trbK* genes are encoded within the *trb* operon, which is tightly regulated by the quorum-sensing regulator TraR and the cognate acylhomoserine lactone (AHL). In the absence of quorum-sensing signals, these proteins are not significantly expressed, and cells harboring such plasmids are efficient Ti plasmid recipients. In the presence of these signals, these strains block the entry of Ti plasmids and instead become efficient conjugal donors.

3.2. Introduction

Many conjugative plasmids are able to block the entry of identical or closely related types of plasmids by creating a functional barrier at the cell surface. This phenomenon is known as entry (or surface) exclusion. Two different types of exclusion determinants are known to cause this phenomenon. Surface-exposed outer membrane proteins, exemplified by TraT of the F plasmid, are thought to block the

[3] Chapter 3 is modified with changes from a published manuscript with the permission from the co-authors and the publisher. Cho, H.*, Pinto, U.M.* and Winans, S.C. (2009) Transsexuality in the rhizosphere: quorum sensing reversibly converts *Agrobacterium tumefaciens* from phenotypically female to male. *J Bacteriol* **191**: 3375-3383. *** U.M.P. and H.C. made equal contributions**. U.M.P. contributed with data for Tables 5, 6, Figs. 1 and 3, and H.C. with data for Tables 3, 4, Figs. 1 and 2.

formation of stable mating aggregates between two donor cells (Sukupolvi and O'Connor, 1990). Other proteins such as TraS of the F plasmid and TrbK of RP4, are located in the inner membrane and inhibit conjugative DNA transfer (Haase *et al.*, 1996; Sukupolvi and O'Connor, 1990).

Entry exclusion of *Agrobacterium* Ti plasmids has not been documented, but it is plausible that they too have such a system (Li *et al.*, 1999). These plasmids are capable of efficient conjugation, and encode a complete suite of conjugative transfer genes, designated *tra* and *trb* genes (Alt-Morbe *et al.*, 1996; Farrand *et al.*, 1996; Li *et al.*, 1999; Zhu *et al.*, 2000). One of these genes, *trbK*, resembles the *trbK* gene of the IncP plasmids RP4, RK2, and R18 (all of which are virtually identical), which mediates entry exclusion of that plasmid (Haase *et al.*, 1995; Haase *et al.*, 1996; Lessl *et al.*, 1991; Lyras *et al.*, 1994). Another Ti plasmid gene, *trbJ*, resembles the *trbJ* gene of RP4, which may or may not contribute to entry exclusion. Lessl et al. and Lyras et al. reported that TrbJ proteins from IncPα plasmids mediate low-level entry exclusion (Lessl *et al.*, 1991; Lyras *et al.*, 1994). Haase et al. presented somewhat conflicting data about the role of TrbJ from RP4 (Haase *et al.*, 1995; Haase *et al.*, 1996). The reasons for these conflicting data are unclear. The *trbJ* and *trbK* genes of RP4 and of Ti plasmids lie with operons of genes that direct mating pair formation (Mpf genes) (Alt-Morbe *et al.*, 1996; Li *et al.*, 1999). The structure encoded by Mpf genes is sometimes referred to as a mating bridge and resembles the family of Type IV systems that are able to translocate DNA and/or protein into foreign cells (Christie *et al.*, 2005). TrbK of RP4 is not required for conjugation (Haase *et al.*, 1995), so its sole function may be in entry exclusion. Similarly, TrbK of pTiC58 is dispensable for conjugation (Li *et al.*, 1999). In contrast, the TrbJ proteins of pTiC58 and of RP4 are essential for conjugation (Haase *et al.*, 1995; Li *et al.*, 1999).

TrbK of RP4 is a lipoprotein that has a lipid attachment motif and is localized mainly to the cytoplasmic membrane (Haase *et al.*, 1996). Its signal sequence is removed proteolytically and one or more acyl groups are added to a cysteine residue at the newly created amino terminus. This cysteine is required for wild type levels of entry exclusion, although residual levels were detectable when this cysteine was altered (Haase *et al.*, 1996). The alteration of the cysteine residue caused a decreased affinity for the cytoplasmic membrane. Significantly, all known Ti plasmid TrbK proteins lack this cysteine residue. They are therefore unlikely to be acylated. Both TrbK and TrbJ proteins are strongly predicted to have cleaved signal sequences (see below), though this prediction has not been experimentally confirmed and the localization patterns of the proteins have not been determined.

All Ti plasmid *tra* and *trb* genes are regulated by the TraR and TraI quorum-sensing system (Fuqua and Winans, 1994) and a variety of plasmids of *Rhizobium*, *Mesorhizobium*, and *Sinorhizobium* spp. regulate conjugation genes in a similar fashion (Gonzalez and Marketon, 2003). TraR resembles the transcription factor LuxR of *Vibrio fischeri*, while TraI resembles the *V. fischeri* LuxI protein, and synthesizes the pheromone 3-oxo-octanoylhomoserine lactone (OOHL). This pheromone binds to and activates TraR. Significantly, both TraR and TraI are encoded on Ti plasmids, and therefore this system detects a quorum of conjugal donors rather than of conjugal recipients. As this system detects only conjugal donors, it seemed plausible that conjugation in *A. tumefaciens* had evolved to occur preferentially between conjugal donors. Although conjugation between donor cells may seem futile, it may have the potentially useful effect of increasing plasmid copy number, as transfer requires conjugative DNA replication. Furthermore, it has been well established that TraR-OOHL complexes increase plasmid copy number by enhancing vegetative replication (Li and Farrand, 2000; Pappas and Winans, 2003).

However, the findings of the present study disproved this hypothesis, as we document that octopine-type and nopaline-type Ti plasmids have entry exclusion systems and that both TrbJ and TrbK can carry out entry exclusion independently and synergistically. In this sense, our findings tend to support the studies of RP4 of the Lessl and Lyras groups (Lessl *et al.*, 1991; Lyras *et al.*, 1994), rather than those of the studies by Haase et al. (Haase *et al.*, 1995; Haase *et al.*, 1996). However, like all *tra* and *trb* genes, *trbJ* and *trbK* are tightly regulated by activated TraR (Fuqua and Winans, 1994; Hwang *et al.*, 1994; Piper *et al.*, 1993), and in the absence of activated TraR, neither TrbJ nor TrbK is significantly expressed and host cells exhibit little or no entry exclusion. These cells, therefore, are efficient recipients, despite the fact that they have Ti plasmids.

3.3. Materials and methods

Strains, oligonucleotides, and reagents

Bacterial strains and plasmids used in this study are described in Table 3.1, while oligonucleotides used for PCR amplification, site-directed mutagenesis and for nuclease S1 protection assays are described in Table 3.2. Antibiotics and ONPG (o-nitrophenyl-β-D-galactopyranoside) were purchased from Sigma-Aldrich. X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was purchased from Gold Biotechnologies. Restriction endonucleases, T4 DNA ligase and T4 polynucleotide kinase were purchased from New England Biolabs. Taq polymerase was purchased from Promega, and [γ-P^{32}]-ATP was purchased from Perkin Elmer.

Quantitative conjugation assays

Conjugative donors and recipients were cultured in AT minimal broth (Tempe *et al.*, 1977) at 27° C for 5 hours, concentrated by centrifugation, combined in a ratio

of 50 recipients per donor, spotted onto AT agar medium, incubated for 2.5 hours for R10-derived donor strain or 18 hours for C58-derived donor strain. Mating was stopped by resuspending the cells from the agar in 1X AT buffer, and then cells were serially diluted and plated on selective AT defined agar medium containing the appropriate antibiotics.

Site directed mutagenesis of *trbJ* and *trbK*

Site-directed mutagenesis of *trbJ* and *trbK* was performed by using a synthetic overlap extension PCR (Sambrook and Russel, 2001). For mutation of *trbJ*, an 1100-bp fragment of pUP404 including a unique *Eco*RI site located upstream of *trbJ* and a unique *Bam*HI site located downstream of *trbJ* was amplified using Platinum Taq Hi Fi DNA Polymerase (Invitrogen). For mutation of *trbK*, a 530-bp fragment of pHC368 including the same restriction sites listed above was amplified. All oligonucleotides used in this study are listed in Table 3.2 and were obtained from Integrated DNA Technologies (Coralville, IA). For *trbJ*, the flanking primers TrbJKF-N and pTacR2 were used in separate reactions with two complementary mutagenic primers, with pUP404 as the template. For *trbK*, the flanking primers were R10TrbKF-N and pTacR1, and the template was pHC368. In both cases, the two PCR products were combined and used as the template in a second round of PCR with the same flanking primers to generate the complete *trbJ* or *trbK* genes. The second set of PCR products was digested with *Eco*RI and *Bam*HI and ligated into pUP404 or pHC368, digested with the same enzymes. These mutations caused a one-codon deletion at the 3'ends of both genes. Mutated sequences were confirmed by automated DNA sequencing.

Table 3.1. Strains and plasmids used in this study.

Strains	Description[a]	reference
WCF5	R10, *traR*, *traR-lacZ*, Km^R	(Fuqua and Winans, 1994)
R10	octopine type strain, pTiR10	S.K. Farrand
C58	nopaline type strain, pTiC58	S.K. Farrand
C58C1RS	Ti plasmid-less derivative of C58, Rif^R, Sm^R	S.K. Farrand
HC158	C58 containing pHC320 inserted into the nopaline-type Ti plasmid pTiC58 by Campbell-type integration, *traR*, *traR-lacZ*, Km^R	This study
HC159 (pYDH902)	R10 cured of pTiR10, and containing cosmid pYDH902	This study
HC161	Strain with polar mutant of *trbD* by the insertion of pHC327	This study
HC162	Strain with polar mutant of *trbJ* by the insertion of pHC328	This study
HC163	Strain with polar mutant of *trbK* by the insertion of pHC329	This study
HC164	Strain with polar mutant of *trbF* by the insertion of pHC330	This study
Plasmids		
pCF218	P*tetR-traR*, Tc^R *rep*-RP4	(Fuqua and Winans, 1994)
pBBR1MCS5	Broad host-range vector, *rep*-pBRR1, Gm^R	(Kovach *et al.*, 1995)
pVIK111	Carries promoterless *lacZ*, ori-R6K, Km^R	(Kalogeraki and Winans, 1997)
pKNG101	*sacB*+, Sm^R, *rep*-R6K, *oriT*-RP4	(Kaniga *et al.*, 1991)
pPR1068	pMALC2 derivative with NdeI site at ATG of *malE*. P*tac*-MBP-lacZα, $lacI^Q$ Ap^R, *ori*-ColE1	Paul Riggs (Peng *et al.*, 2001)
pPZP200 and pPZP201	broad host-range vector, *rep*-pVS1, Sp^R	(Hajdukiewicz *et al.*, 1994)
pJZ335	*traR* from pTiA6NC cloned into pPZP201	(Zhu and Winans, 2001)
pJZ381	EcoRI fragment containing *traR* cloned in pBBR1MCS5	(Cho and Winans, 2007)
pYDH902	Cosmid containing *rep* and *traI-trb* operons, *rep*-RP4, Tc^R	(Dessaux *et al.*, 1987)

Table 3.1. (continued)

pHC011	pPR1068 digested with NdeI and SacI, with replacement by a linker containing NdeI-KpnI-SacI sites to delete the *malE* gene. P*tac* is fused to NdeI-KpnI-SacI-AvaI-XmnI-EcoRI-BamHI-XbaI-SalI-PstI-HindIII. *rep*-ColE1, Ap^R	This study
pHC012	pHC011 digested with EcoRV and KpnI, ligated to pBBRMCS5 after digestion with SphI and KpnI, with 3'-end fill-in of SphI site with Klenow fragment of DNA polymerase I. P*tac* is fused to NdeI-KpnI-ApaI-XhoI-SalI-Bsp106I-ClaI-HindIII-EcoRI-PstI-SmaI-BamHI-SpeI-XbaI-BstXI-SacI. *rep*-pBBR1, Gm^R	This study
pHC320	pVIK111 containing an EcoRI-XbaI fragment containing the 5' end of *traR* and upstream sequences; *rep*-R6K, Km^R	This study
pHC327	PCR fragment of *trbD* made using oligos TDF and TDR, and cloned into pKNG101 for Campbell recombination mutagenesis. *rep*-R6K, Sm^R	This study
pHC328	PCR fragment of *trbJ* made using oligos TJF and TJR, and cloned into pKNG101 for Campbell recombination mutagenesis. *rep*-R6K, Sm^R	This study
pHC329	PCR fragment of *trbK* made using oligo TKF and TKR, and cloned into pKNG101 for Campbell recombination mutagenesis. *rep*-R6K, Sm^R	This study
pHC330	PCR fragment of *trbF* made using oligo TFF and TFR, and cloned into pKNG101 for Campbell recombination mutagenesis. *rep*-R6K, Sm^R	This study
pHC335	pJZ335 digested with BamHI and ligated, removing a small BamHI fragment between P*lac* and *traR*. *traR* cloned into pPZP201 *rep*-pVS1, Sp^R	This study
pHC361	PCR fragment containing *trbJK* made using oligos trbJKF-N and R10 trbKJK3, digested with BamHI, and cloned into pHC012; P*tac*-*trbJK*. *rep*-pBBR1; Gm^R	This study
pHC364	*trbBCDEJK* cloned as a HindIII-BamHI fragment into pHC012, P*tac*-*trbBCDEJK* *rep*-pBBR1, Gm^R	This study

Table 3.1. (continued)

pHC368	PCR fragment containing *trbK*, made using oligos R10-trbKF-N and R10 trbKJK3, digested with BamHI, and cloned into pHC012; P*tac*-*trbK*, *rep*-pBBR1, Gm^R	This study
pUP200	1236-nucleotide DNA fragment made by PCR amplification using pHC012 as template and oligos MfeI-For and NsiI-Rev as primers, cloned into EcoRI-PstI gap of pPZP200, with the P*tac* promoter upstream of multiple cloning site of pHC012. *rep*-pVS1, Sp^R	This study
pUP402	PCR fragment containing *trbJ* made using oligos trbJKF-N and trbJR-N and cloned into pHC012; P*tac*-*trbJ* *rep*-pVS1, Gm^R	This study
pUP403	PCR fragment containing *trbJ* made using oligos trbJKF-N and trbJR-N and cloned into pUP200; P*tac*-*trbJ*. *rep*-pVS1, Sp^R	This study
pUP404	*trbK* from pHC368 cloned into pUP200. P*tac*-*trbK*, *rep*-pVS1, Sp^R	This study
pUP405	Derivative of pHC368 lacking the 3' codon of *trbK*	This study
pUP406	Derivative of pUP404 lacking the 3' codon of *trbK*	This study
pUP407	Derivative of pUP402 lacking the 3' codon of *trbJ*	This study
pUP408	Derivative of pUP403 lacking the 3' codon of *trbJ*	This study

Nuclease S1 protection assays

RNA was isolated from cells cultured to late log phase and harvested in the presence of two volumes of RNAprotect Bacteria Reagent (Qiagen) per volume of culture. Cell pellets were frozen at -80° C. Lysozyme (200 µL of 10 mg/ml solution) and 700 µL of buffer RLT (Qiagen) were added to the frozen cell pellets, and tubes were subjected to a vigorous vortex. Lysates were clarified by centrifugation for 2 min, and RNA was precipitated from the supernatant by addition of 500 µL ethanol. Samples were applied to RNeasy spin columns (Qiagen) and centrifuged for 15 s at 10,000 rpm. Buffer RW1 (350 µL) was added to each column, and columns were centrifuged for 15 s at 10,000 rpm. DNase I was diluted eightfold in RDD buffer (Qiagen), and 80 µL per column were added. After 15 m incubation, columns were washed successively with 350 µL of buffer RW1 and 500 µL of buffer RPE, and RNA was eluted using 40 µL of Rnase-free water.

Oligonucleotides were radiolabelled with [λ-P^{32}]ATP and T4 DNA kinase. A 500 pg sample of radiolabelled oligonucleotides was hybridized with 20 µg of total RNA for 10 h at 42°C and then digested with 250 units of nuclease S1 for 1 h at 37°C. Reaction mixtures were then ethanol precipitated and suspended in 5 µl of 0.1 M NaOH, and 5 µl of formamide loading dye was added. Five microliters of each sample were size-fractionated using 18% denaturing Tris-borate-EDTA polyacrylamide gels and quantified using a Storm phosphorImager (Model 840, Molecular Dynamics). A 2.5 pg aliquot of ^{32}P-labelled nondigested oligonucleotide was added to one lane of each gel.

Table 3.2. Oligonucleotides used in this study.

Oligonucleotides	Sequence[a]
TDF	5'-GCT*GGATCC*ATCGCTGGTGCGATGCTG-3'
TDR	5'-GCT*TCTAGA*ACCATCTCGACCCCTTCAG-3'
TJF	5'-GCT*GGATCC*AATGGGCAATGTCGAAGATG-3'
TJR	5'-GCT*TCTAGA*GCGCGAGAATGACGATCAG-3'
TKF	5'-GCT*TCTAGA*GCAGGCACAAAAGGATCTG-3'
TKR	5'-CGGCGATACCGACCTCGATG-3'
TFF	5'-GCT*GGATCC*TCATCCCCTACATCGTTGAG-3'
TFR	5'-GCT*TCTAGA*CCATGCCTTTCAAAGCTGTG-3'
Oligonucleotides used to subclone *trb* genes	
TrbJKF-N	5'-GCT*GAATTC*GCAAAGGGGGATCGCCCATG-3'
TrbJR-N	5'- GTC*GGATCC*GAGAATGACGATCAGACGCG -3'
R10-trbKF-N	5'-GCT*GAATTC*GACGATGGAGCCACGCTGGTG-3'
R10-trbKJK3	5'-ATGAACATGATGCGTTTGAC-3'
Oligonucleotides used to clone P*tac-lacZ* on pPZP200	
MfeI-For	5'-GTC*CAATTG*TATACGCAAGGCGACAAGGTG-3'
NsiI-Rev	5'-GTC*ATGCAT*ACTTATTCAGGCGTAGCACCA-3'
Oligonucleotites used for mutagenesis	
pJ_W269Stop-F	5'- GAGCCACGCTGATGAGCTCGC - 3'
pJ_W269Stop-R	5'- GCGAGCTCATCAGCGTGGCTC - 3'
pK_W75Stop-F:	5'- GAAACCGAGATGATAGTTCACC - 3'
pK_W75Stop-R	5'- GGTGAACTATCATCTCGGTTTC - 3'
pTacR1	5'- ACGACGTTGTAAAACGACGGC - 3'
pTacR2	5'- GCCATTCAGGCTGCGCAACTG - 3'
Oligonucleotide used for Nuclease S1 protection assay	
trbKS1	5'-GGCTGGACAGTAATCCAGGTGCCGATGCCTGCACTACGAC-3'
23SRNAS1	5'-AGGCTCGGGCTCCGACTGTTTGTAGGCATCCGGTTTCAG-3'

[a] Italics indicate restriction endonucleases cleavage sites used in plasmid construction.

3.4. Results

Two Ti plasmids encode functional and tightly regulated entry exclusion systems

The overexpression of TraR in strains containing the native *traI* gene causes constitutive expression of all genes of the quorum-sensing regulon (Fuqua and Winans, 1994). We reasoned that any entry exclusion gene might also be regulated by TraR, and if so, would most likely be expressed constitutively in strains overexpressing TraR. The overexpression of TraR during conjugation also relieves the requirement for octopine, which is otherwise needed to induce the transcription of the native *traR* gene (Fuqua and Winans, 1994) and therefore tends to make conjugation data more reproducible. We measured the efficiency of Ti plasmid transfer from R10 derivative WCF5(pJZ381), which overexpresses TraR (Table 3.1), to two recipient strains: R10(pHC335), which also overexpresses TraR, and R10(pPZP201), which does not. Both recipient strains carried the Ti plasmid pTiR10, which is virtually identical to other so-called octopine-type Ti plasmids (Zhu *et al.*, 2000). The former strain gave rise to 300-fold fewer transconjugants than the latter strain (Table 3.3, first two line), indicating that either TraR itself or, more likely, the product of a TraR-regulated gene mediated a robust level of entry exclusion.

Table 3.3. Entry exclusion of octopine-type and nopaline-type Ti plasmids by homologous and heterologous recipients.

Donor	Recipient	Relevant protein(s) expressed in recipient	No. of transconjugants per Donor (SD)[a]	Exclusion coefficient[b]
Homologous Recipients				
WCF5(pJZ381)[c]	R10(pPZP201)	none	0.94 (0.2)	(1)
WCF5(pJZ381)	R10(pHC335)	TraR, TraI, TraA-H, TrbB-L	0.003 (0.001)	313
HC158(pJZ381)[d]	C58(pPZP201)	None	0.22 (0.03)	(1)
HC158(pJZ381)	C58(pHC335)	TraR, TraI, TraA-H, TrbB-L	0.0008 (0.0004)	275
Heterologous Recipients				
WCF5(pJZ381)	C58(pPZP201)	none	0.032 (0.01)	(1)
WCF5(pJZ381)	C58(pHC335)	TraR, TraI, TraA-H, TrbB-L	0.005 (0.002)	64
HC158(pJZ381)	R10(pPZP201)	none	0.001 (0.001)	(1)
HC158(pJZ381)	R10(pHC335)	TraR, TraI, TraA-H, TrbB-L	< 0.00001[e]	> 100
Recipients lacking a Ti plasmid				
WCF5(pJZ381)	C58C1RS(pPZP201)	none	0.47 (0.22)	(1)
WCF5(pJZ381)	C58C1RS(pHC335)	TraR	0.55 (0.08)	0.85
HC158(pJZ381)	C58C1RS(pPZP201)	none	0.35 (0.035)	(1)
HC158(pJZ381)	C58C1RS(pHC335)	TraR	0.42 (0.13)	0.8

[a] Transconjugants were selected using the Km^r gene of the Ti plasmid and the Sp^r gene of pPZP201 or pHC335. In mock conjugations, we did not detect spontaneous resistance to either kanamcin or spectinomycin. The data are the averages of results from three independent experiments, with the standard deviations shown in parentheses.

[b] The exclusion coefficient is the number of transconjugants per donor for the no-exclusion control (lines with exclusion coefficient of 1) divided by the number of transconjugants per tested recipient strains per donor.

[c] R10-derived strain containing a Km^r gene on the octopine-type Ti plasmid.

[d] C58-derived strain containing a Km^r gene on the nopaline-type Ti plasmid.

[e] No transconjugants were detected in an assay mixture containing 100,000 donor bacteria.

Similar experiments were carried out using strains carrying the nopaline-type Ti plasmid pTiC58. Here, strain HC158(pJZ381) was used as a Ti plasmid donor. This strain contains a nopaline-type Ti plasmid that has a kanamycin resistance gene to facilitate the selection of transconjugants. This strain also overexpresses TraR from pJZ381. Strains C58(pHC335) and C58(pPZP201) were used as recipients. The former recipient yielded approximately 300-fold fewer transconjugants than the latter (Table 3.3, third and fourth lines). We conclude that TraR or a TraR-regulated gene in pTiC58 can exclude conjugal entry of the same type of plasmid.

We also tested the abilities of nopaline-type Ti plasmids to exclude octopine-type Ti plasmids, and vice versa. The octopine-type Ti plasmid present in WCF5(pJZ381) conjugated approximately 60 fold less efficiently into a strain containing a nopaline-type Ti plasmid and expressing TraR than into a congenic strain not expressing TraR [Table 3.3, lines for heterologous recipients C58(pPZP201) and C50(pHC335)]. Similar results were obtained with the reciprocal cross (Table 3.3, last two lines for heterologous recipients). In the first cross, entry exclusion appeared to be slightly weaker than in either homologous crosses [Table 3.3, compare line for heterologous recipient C58(pHC335) with lines for homologous recipients R10(pHC335) and C58(pHC335)], suggesting that entry exclusion determinants of the nopaline-type Ti plasmid may function more effectively in blocking a homologous donor than in blocking a heterologous one. For the second cross [Table 3.3, line for heterologous recipient R10(pHC335)], no such conclusion is possible. No transconjugant colonies were detected, suggesting very strong entry exclusion. However, relatively few transconjugants were detected with the negative control [Table 3.3, line for heterologous recipient R10(pPZP201)], suggesting either that a TraR-independent entry exclusion acted in the recipient to block entry or that the plasmid from the donor conjugated inefficiently into this recipient.

To confirm that TraR mediates entry exclusion indirectly, we measured conjugation using recipient strains lacking Ti plasmids. Strains C58C1RS(pHC335) and C58C1RS(pPZP201) lack Ti plasmids, and the former strain expresses TraR while the latter one does not. Neither strain excluded the entry of either Ti plasmid (Table 3.3, lines for recipients lacking a Ti plasmid), indicating that TraR is not sufficient for entry exclusion and that it functions by activating one or more entry exclusion genes.

These data also allow us to compare a strain containing a Ti plasmid but lacking TraR with a strain lacking a Ti plasmid. Strains C58(pPZP201) and C58C1RS(pPZP201) are identical except for the presence or absence of a Ti plasmid. Neither strain overexpresses TraR. These two strains showed little if any difference in their inability to exclude either Ti plasmid (Table 3.3, compare the first line with the fourth-to-last line and the third line with second-to-last line). This finding indicates that entry exclusion determinants are not significantly expressed in the absence of active TraR.

Identification of the entry exclusion determinants encoded in the Ti plasmid

As described above, plasmid RP4 has a *trb* operon that resembles those of Ti plasmids (Fig. 3.1). Within the RP4 operon, the *trbK* gene encodes a product required for entry exclusion (Haase *et al.*, 1995; Haase *et al.*, 1996; Lessl *et al.*, 1991; Lyras *et al.*, 1994), while the *trbJ* product may (Lessl *et al.*, 1991; Lyras *et al.*, 1994) or may not (Haase *et al.*, 1995; Haase *et al.*, 1996) play an accessory role. TrbK of RP4 is 23.5% and 18.2% identical to the TrbK proteins of octopine- and nopaline-type Ti plasmids, respectively, while TrbJ of RP4 is 20.7% identical to both Ti plasmid TrbJ proteins. TrbK proteins of Ti plasmids lack the acylation site of TrbK of RP4, suggesting that they may be nonfunctional or weakly functional. Both TrbJ and TrbK were strongly predicted by program SignalP-HMM to have cleaved signal sequences

(probability, 1.0). The cleavage of TrbJ was predicted to remove 33 residues, while the cleavage of TrbK was predicted to remove 21 residues.

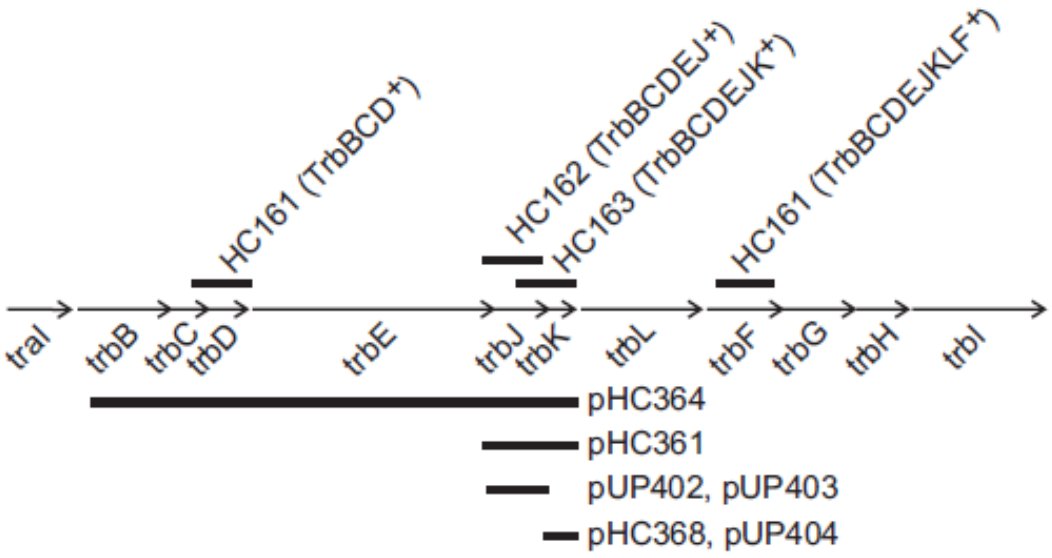


Figure 3.1. Genetic organization of the *trb* operon of an octopine-type Ti plasmid. Short thick lines above the genetic map represent DNA fragments that were used in suicide plasmids to create transcriptionally polar mutations upon Campbell-type integration. Fragments of the *trb* region overexpressed by fusion to the P*tac* promoter are shown beneath the genetic map.

We sought to determine whether TrbK and/or TrbJ of an octopine-type Ti plasmid plays a role in entry exclusion. To address this question, we compared strain R10(pHC335), which contains a native octopine-type Ti plasmid and overexpresses TraR, with strain HC159(pYDH902)(pCH335), which lacks the Ti plasmid, overexpresses TraR, and contains a cosmid (pYDH902) that carries the *trb* and *rep* operons (Dessaux *et al.*, 1987). The donor strain in this experiment was WCF5(pJZ381). Both recipient strains exhibited entry exclusion, and in both cases, TraR overexpression was required (Table 3.4, first four lines). A similar strain,

HC159(pYDH902)(pPZP201), which does not express TraR, showed a low but detectable level of exclusion [Table 3.4, lines for R10(pPZP201) and HC159(pYDH902)(pPZP201)], due possibly to elevated basal expression of entry exclusion determinants from the multicopy plasmid pYDH902. These data indicate that all genes essential for entry exclusion lie within pYDH902 and probably within the *trb* operon.

To more closely localize the genes responsible for entry exclusion, we constructed four insertion mutations in the *trb* operon that are predicted to exert strong transcriptional polarity on downstream genes. We used derivatives of the suicide plasmid pKNG101 containing various *trb* fragments. The insertion in HC161 expresses TrbB, TrbC, and TrbD but not TrbE, TrbJ, TrbK, TrbL, TrbF, TrbG, TrbH, or TrbI (Fig. 3.1). This mutation blocked virtually all entry exclusion [Table 4, line for HC161(pHC335)]. Strain HC162(pHC335) expresses TrbB, TrbC, TrbD, TrbE, and TrbJ and showed approximately fourfold fewer transconjugants than the negative control (Table 3.4), suggesting a role of TrbJ and/or TrbE in entry exclusion. Strain HC163(pHC335), which expresses TrbB, TrbC, TrbD, TrbE, TrbJ, and TrbK (Fig. 3.1), strongly expressed entry exclusion (Table 3.4), indicating that TrbK plays a major role. This strain expressed entry exclusion levels similar to HC164(pHC335), which expresses two additional Trb proteins, and to R10(pHC335), which expresses all Trb proteins (Table 3.4), suggesting that the genes downstream of *trbK* do not have any role in entry exclusion.

Table 3.4. Mapping the entry exclusion locus of the Ti plasmid by using polar insertion mutations within the *trb* operon.

Recipient[a]	Relevant Proteins Expressed in Recipient	Conjugation efficiency[b] (SD)	Exclusion Coefficient[c]
Recipients without *trb* mutations			
R10(pPZP201)	none (vector control)	0.90 (0.1)	(1)
R10(pHC335)	TraR, TraI, TraA to TraH, TrbB to TrbL	0.006 (0.0007)	150
HC159(pYDH902) (pPZP201)	None	0.11 (0.015)	8.2
HC159(pYDH902) (pHC335)	TraR, TraI, Trb	0.0005 (0.00008)	1800
Recipients with *trb* genes mutated using transcriptionally polar insertion mutations			
HC161(pHC335)	TraR, TrbBCD	0.84 (0.1)	1.07
HC162(pHC335)	TraR, TrbBCDEJ	0.19 (0.02)	4.7
HC163(pHC335)	TraR, TrbBCDEJK	0.007 (0.001)	129
HC164(pHC335)	TraR, TrbBCDEJKLF	0.007 (0.001)	129
Recipients with Trb proteins expressed from a multicopy plasmids via a *tac* promoter			
R10(pHC012)	none (vector control)	0.18 (0.02)	(1)
R10(pJZ381)	TraR, TraI, TraA-H, TrbB-L	0.0004 (0.0002)	450
R10(pHC364)	TrbBCDEJK	0.008 (0.004)	22.5
R10(pHC361)	TrbJK	0.0012 (0.0005)	150
R10(pHC368)	TrbK	0.025 (0.006)	7.2

[a] The donor strain in each experiment was WCF5(pJZ358), which overexpresses TraR and has an octopine-type Ti plasmid marked with a Km^r determinant (Fuqua and Winans, 1994). Transconjugants were selected using the Km^r gene of the Ti plasmid and the Sp^r gene of pPZP201 or pHC335 (first eight lines) or using the Gm^r gene of pHC012 and its derivatives (last four lines).

[b] number of transconjugants per donor.

[c] The exclusion coefficient is the number of transconjugants per donor for the no-exclusion control (lines with exclusion coefficient of 1) divided by the number of transconjugants of the tested recipient strain per donor.

We also tested the expression of Trb proteins expressed from a Ptac promoter of a multicopy plasmid. Plasmid pHC364 expresses TrbB, TrbC, TrbD, TrbE, TrbJ, and TrbK (Fig. 3.1), and expressed entry exclusion, albeit at a reduced level [Table 3.4, lines for R10(pJZ381) and R10(pHC364)]. Plasmid pHC361, which expresses only TrbJ and TrbK (Fig. 3.1), expressed high levels of entry exclusion, while plasmid pHC368, which expresses only TrbK, expressed a low level of entry exclusion (Table 3.4, last two lines).

To further measure the effects of TrbJ and TrbK on entry exclusion, we expressed these proteins using separate, compatible plasmids in recipient strains. We made a series of fusions using plasmids pHC012 and pUP200, both of which have Ptac promoters and *lacZ*α genes. P*tac-trbJ* fusions were constructed in such a way that the *lacZ*α gene was translationally fused to the stop codon of *trbE* (which lies immediately upstream of *trbJ* in the native Ti plasmid). This was done to mimic any possible translational coupling betweeen *trbE* and *trbJ*. Similarly, P*tac-trbK* fusions were made in such a way that the *lacZ*α gene was translationally fused to the stop codon of *trbJ*.

Expressing TrbJ alone from a derivative of pBBRMCS5 (pUP402) decreased conjugation approximately ninefold, while expressing it from a derivative of pPZP200 (pUP403) caused a fourfold decrease (Table 3.5, third and fourth lines). This difference is most likely attributable to a difference in copy number, as the Ptac-*trbJ* fusions of the two plasmids are identical in sequence. The expression of TrbK alone in these two vectors caused similar decreases in conjugation (Table 3.5, fifth and sixth lines). Most importantly, coexpressing these two proteins from compatible plasmids caused a strong additional decrease in conjugation (Table 3.5, seventh and eighth lines). We conclude that TrbJ and TrbK make independent contributions to entry exclusion and that the presence of both proteins has a synergistic effect.

Table 3.5. Expression of TrbJ and TrbK in recipient using multicopy plasmids[a]

pBBRMCS5 derivative (description or genotype)	pPZP200 derivative (description or genotype)	Conjugation efficiency[b] (SD)	Exclusion coefficient[c]
pHC012 (vector)	pUP200 (vector)	0.69 (0.2)	(1)
pJZ381 (*traR*)[d]	pPZP200 (vector)	0.002 (0.001)	345
pUP402 (P*tac-trbJ*)	pUP200 (vector)	0.08 (0.009)	8.6
pHC012 (vector)	pUP403 (P*tac-trbJ*)	0.18 (0.03)	3.8
pHC368 (P*tac-trbK*)	pUP200 (vector)	0.10 (0.04)	6.9
pHC012 (vector)	pUP404 (P*tac-trbK*)	0.21 (0.04)	3.3
pUP402 (P*tac-trbJ*)	pUP404 (P*tac-trbK*)	0.007 (0.002)	98.6
pHC368 (P*tac-trbK*)	pUP403 (P*tac-trbJ*)	0.022 (0.004)	31.4
pHC361 (P*tac-trbJK*)	pPZP200 (vector)	0.002 (0.0008)	345

[a] The donor strain in each experiment was WCF5(pCF218), which overexpresses TraR and has an octopine-type Ti plasmid marked with a Km^r determinant. Transconjugants were selected using the Km^r gene of the Ti plasmid, the Sp^r gene of pUP200 or its derivatives, and the Gm^r gene of pHC012 or its derivatives.
[b] Number of transconjugants per donor.
[c] The exclusion Coefficient is the number of transconjugants per donor for the no-exclusion control (top line) divided by the number of transconjugants of the tested recipient strain per donor.
[d] The overexpression of TraR by pJZ381induces expression of all *tra* and *trb* genes (Fuqua and Winans, 1994).

Interestingly, a strain expressing TrbJ and TrbK from separate plasmids showed less entry exclusion than a strain expressing these proteins on a single plasmid (Table 3.5, last three lines). To ensure that TrbK was expressed at similar levels in these strains, we assayed for accumulation of TrbK mRNA. Plasmid pHC368, which has a P*tac-trbK* fusion, expressed considerably more TrbK mRNA than pHC361, which has a P*tac- trbJK* fusion (Fig. 3.2). Despite this result, the former plasmid expressed entry exclusion more weakly than the latter plasmid. This underscores the importance of TrbJ in this process and supports the conclusion that these proteins preferentially act *in cis* or together.

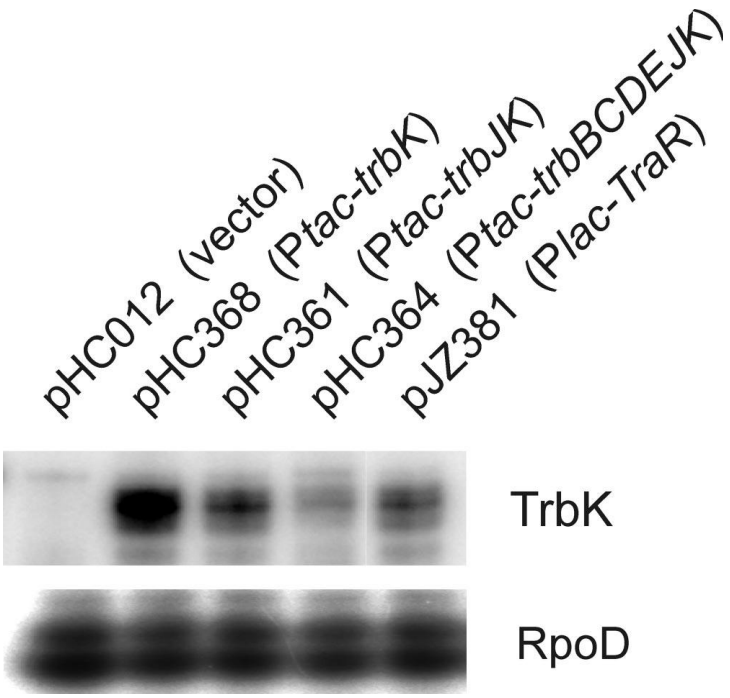


Figure 3.2. Nuclease S1 protection assays. Results show *trbK* transcript levels in recipients containing fusions between Ptac and the indicated *trb* genes (top) and *rpoD* transcript levels for each strain (bottom). All strains are derivatives of *A. tumefaciens* strain R10, which contains pTiR10. Plasmid pJZ381 expresses the *trbK* gene on the Ti plasmid.

In the course of searching for proteins homologous to TrbK, we fortuitously noticed sequence similarity between TrbK and TrbJ. The C-terminal 12 amino acid residues of these proteins are identical or similar (Fig. 3.3). This similarity is found among a variety of plasmids of *Agrobacterium*, *Rhizobium*, and *Sinorhizobium* and in two plasmids found in *Nitrobacter hamburgensis* and *Oligotropha carboxidovorans*, (Fig. 3.3). The latter two bacteria express TrbJ and TrbK proteins that are strongly similar to those of *A. tumefaciens* and its close relatives. The last aminoacid residue of the protein shows remarkable conservation, even in more distantly related proteins from the IncP-type plasmids. A small number of other cognate TrbJ and TrbK proteins also show sequence similarities at their C termini, but the majority do not (Fig. 3.3 and data not shown).

Since the last amino acid residue of TrbJ and TrbK, both tryptophans, are completely conserved not only in all Ti plasmid proteins but also in more distantly related proteins from other plasmids, it seemed plausible that these residues may play a crucial role in protein function. To test the functional importance of the similar C-termini of TrbK and TrbJ, we deleted the last amino acid residue by using site-directed mutagenesis and tested the mutated proteins for their role in entry exclusion. A truncated TrbK protein had virtually null phenotype when expressed alone and had little if any synergistic effect when coexpressed with TrbJ (Table 3.6). The corresponding mutation on TrbJ also had a strong impact on the ability of the protein to mediate entry exclusion, though the mutant TrbJ still mediate a low level of exclusion when expressed together with wild-type TrbK (Table 3.6, 2^{nd}, 3^{rd}, 6^{th}, 7^{th}, 11^{th}, and 12^{th} lines). When both mutant proteins were expressed together in the cell, entry exclusion was negligible compared to that mediated by wild-type proteins (Table 6, last four lines). Overall, these results confirm the prediction that the C-termini of the two proteins play a crucial role in entry exclusion.

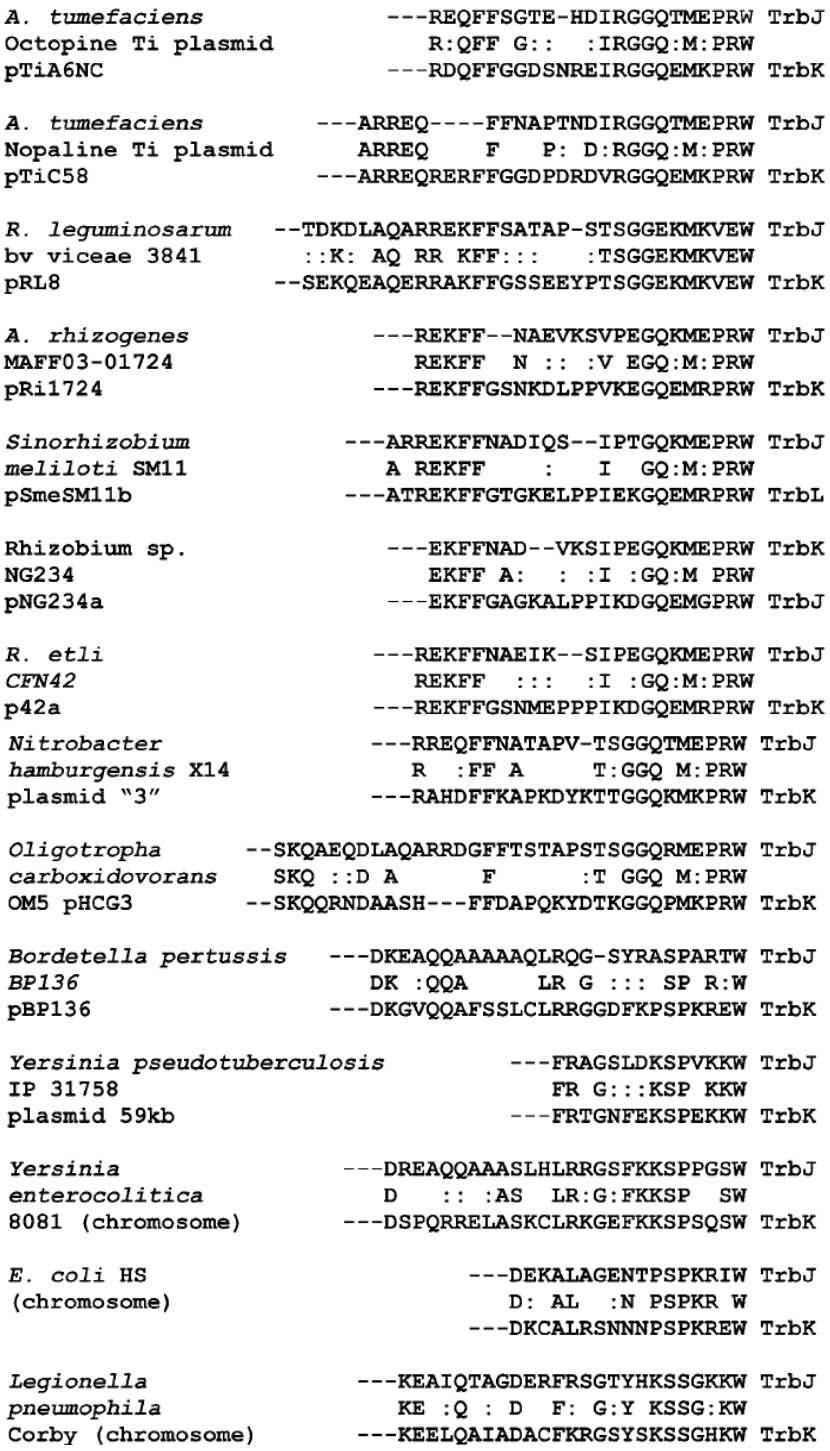


Figure 3.3. Alignment of the C-termini of TrbJ and TrbK proteins of selected conjugation systems. Sequence similarities between TrbJ and TrbK pairs were obtained using the Megalign program (DNASTAR). Colons indicate conservative substitutions. *R. leguminosarum, Rhizobium leguminosarum*; *A. rhizogenes, Agrobacterium rhizogenes, R. etli, Rhizobium etli; E. coli, Escherichia coli*.

Table 3.6. Effects of deleting the C-terminal Trp residue of TrbJ and TrbK[a]

pBBRMCS5 derivative (description or genotype)	pPZP200 derivative (description or genotype)	Conjugation Efficiency[b](SD)	Exclusion coefficient
pHC012 (vector)	pUP200 (vector)	2.62 (0.88)	1
pHC012 (vector)	pUP403 (P*tac-trbJ*)	0.81 (0.68)	3.2
pHC012 (vector)	pUP408 (P*tac-trbJ**)	1.66 (0.51)	1.6
pHC012 (vector)	pUP404 (P*tac-trbK*)	0.26 (0.06)	10.0
pHC012 (vector)	pUP406 (P*tac-trbK**)	1.55 (0.69)	1.7
pUP402 (P*tac-trbJ*)	pUP200 (vector)	0.35 (0.23)	7.4
pUP407 (P*tac-trbJ**)	pUP200 (vector)	0.63 (0.22)	4.2
pHC368 (P*tac-trbK*)	pUP200 (vector)	0.31 (0.14)	8.5
pUP405 (P*tac-trbK**)	pUP200 (vector)	1.58 (0.68)	1.7
pUP402 (P*tac-trbJ*)	pUP406 (P*tac-trbK**)	0.17 (0.07)	15.2
pUP407 (P*tac-trbJ**)	pUP404 (P*tac-trbK*)	0.09 (0.04)	28.4
pHC368 (P*tac-trbK*)	pUP408 (P*tac-trbJ**)	0.21 (0.08)	12.7
pUP405 (P*tac-trbK**)	pUP403 (P*tac-trbJ*)	0.44 (0.12)	5.9
pUP407 (P*tac-trbJ**)	pUP406 (P*tac-trbK**)	0.52 (0.26)	5.1
pUP405 (P*tac-trbK**)	pUP408 (P*tac-trbJ**)	1.36 (0.07)	1.9
pUP402 (P*tac-trbJ*)	pUP404 (P*tac-trbK*)	0.018 (0.008)	143.9
pHC368 (P*tac-trbK*)	pUP403 (P*tac-trbJ*)	0.055 (0.018)	47.6

[a] The donor strain in each experiment was WCF5(pCF218), which overexpresses TraR and has an octopine-type Ti plasmid marked with a Km^r determinant. Transconjugants were selected using the Km^r gene of the Ti plasmid, the Sp^R gene of pUP200 or its derivatives, and the Gm^r gene of pHC012 or its derivatives. The symbol * denotes a deletion of the last residue of the corresponding protein.
[b] Number of transconjugants per donor.

3.5. Discussion

It is well established that strains lacking active TraR do not conjugate or conjugate at extremely low levels (Fuqua and Winans, 1994; Hwang *et al.*, 1994; Piper *et al.*, 1993). We now show that such strains also do not express entry exclusion functions and therefore readily act as conjugative recipients of Ti plasmids. The fact that C58(pPZP201) and C58C1(pPZP201) were virtually identical in their abilities to receive a Ti plasmid, even though the former has a Ti plasmid while the latter lacks it, indicates that entry exclusion genes are tightly regulated. Strains containing conjugal plasmids but not expressing conjugation or entry exclusion functions are sometimes

referred to as "female phenocopies" (Press *et al.*, 1971). Female phenocopies are generally detected after long term culturing of a strain at stationary phase. In the case of *A. tumefaciens*, cultures that do not express active TraR are female phenocopies, even when actively growing, a consequence of the extremely tight regulation of the *tra-trb* regulon.

The finding that TrbJ and TrbK mediate entry exclusion was initially surprising. On the one hand, TrbJ and TrbK of RP4 have been described as mediating this property. However, as described above, there is considerable controversy about the role of TrbJ (Haase *et al.*, 1995; Haase *et al.*, 1996; Lessl *et al.*, 1991; Lyras *et al.*, 1994). Furthermore, *A. tumefaciens* TrbK lacks a cysteine residue that is critical for normal function of the RP4 protein, suggesting that *trbK* of *A. tumefaciens* may be a pseudogene. It seemed plausible that *A. tumefaciens* might not exhibit entry exclusion, as described above. Finally, it seemed counterintuitive for entry exclusion functions to be located within a tightly regulated operon. One may imagine a priori that exclusion genes may be needed even when the Tra/Trb regulon is not expressed, and it would seem a simple evolutionary step for these genes to be expressed constitutively.

As described above, pHC368, which expresses just TrbK, makes considerably more TrbK mRNA than pHC361, which expresses TrbJ and TrbK (Fig. 3.2). Despite this, the former plasmid expresses the entry exclusion phenotype more weakly than the latter. This finding highlights the importance of TrbJ in entry exclusion. However, pHC368 expresses entry exclusion more weakly than pHC361 even in the presence of a second plasmid expressing TrbJ (Table 3.5, lines for pHC368 and pHC361). The P*tac-trbJ* fusions of pHC361, pUP402, and pUP403 are identical, making it unlikely that TrbJ is expressed at greatly different levels in these three plasmids. The most likely interpretation is that TrbJ and TrbK function more effectively when expressed

in cis than *in trans*. An alternative interpretation is that the TrbJ and TrbK interact and do so more effectively if expressed at the same location.

As noted earlier, we found a curious sequence similarity between the C-termini of TrbJ and TrbK. Mature TrbK proteins are predicted to be quite small, approximately 50 amino acid residues in length, and the C-terminal 15 residues therefore constitute a rather large fraction of the entire protein. The C-termini of TrbK proteins are also far more conserved than other parts of these proteins (data not shown), suggesting that these residues may be crucial for protein function. In some cases, a TrbK protein of one plasmid may resemble TrbJ from the same plasmid more strongly than it resembles TrbK proteins of other plasmids (Fig. 3.3). This pattern suggests that the TrbJ and TrbK proteins encoded by a particular plasmid may coevolve by a process resembling gene conversion. In light of the overlapping functions of TrbJ and TrbK, it seemed tempting to speculate that the C-termini of both proteins may play a crucial role in entry exclusion. In fact, the results of deleting the last amino acid residues of both proteins confirmed this hypothesis (Table 3.6). It may be noteworthy that the C-terminal five amino acid residues of TrbK of RP4 are essential for activity (Haase *et al.*, 1996). Interestingly, our results show that TrbK protein cannot tolerate a truncation eliminating its last amino acid residue. However, TrbJ can still function in the presence of TrbK, albeit rather poorly. These results also suggest that these two proteins might interact for proper function, but this hypothesis remains to be tested.

The finding that a bacterium having a Ti plasmid but not expressing the Tra-Trb regulon is a female phenocopy may imply interesting ecological consequences. For example, one could imagine a situation in which two strains of *A. tumefaciens*, one containing an octopine-type Ti plasmid similar to pTiA6 and the other having a nopaline-type Ti plasmid similar to pTiC58, colonize the same crown gall tumor. One

could imagine furthermore that there is an abundance of octopine but very little or no agrocinopine A or B (the conjugal opines for pTiC58). Conjugal opines are required for conjugation, as conjugal opines are required for transcription of both *traR* genes (Fuqua and Winans, 1994; Piper *et al.*, 1993). In such a scenario, the octopine-type Ti plasmid would both conjugate and block the entry of nopaline-type Ti plasmid, while the nopaline-type Ti plasmid would do neither. If an octopine-type Ti plasmid conjugated into a strain already containing a nopaline-type Ti plasmid, the transconjugants would contain both Ti plasmids. These plasmids are incompatible at the level of DNA replication and would segregate into different daughter cells upon cell division. As a result, new combinations of host strains and Ti plasmids may appear. Thus, an active entry exclusion system would prevent futile transfer of Ti plasmids between identical strains, but would allow reassortment of Ti plasmids and heterologous host strains even if those strains already contained heterologous Ti plasmids.

Acknowledgments

We thank the members of our laboratory for helpful discussions and for critical reviewing of this text. This work was supported by a grant from the National Institute of General Medical Sciences (GM42893). U.M.P acknowledges the financial support of the Brazilian government through a fellowship grant from the "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

REFERENCES

Alt-Morbe, J., Stryker, J.L., Fuqua, C., Li, P.L., Farrand, S.K., and Winans, S.C. (1996) The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. *J Bacteriol* **178**: 4248-4257.

Cho, H., and Winans, S.C. (2007) TraA, TraC and TraD autorepress two divergent quorum-regulated promoters near the transfer origin of the Ti plasmid of *Agrobacterium tumefaciens*. *Mol Microbiol* **63**: 1769-1782.

Christie, P.J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S., and Cascales, E. (2005) Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol* **59**: 451-485.

Dessaux, Y., Tempe, J., and Farrand, S.K. (1987) Genetic analysis of mannityl opine catabolism in octopine-type *Agrobacterium tumefaciens* strain 15955. *Mol Gen Genet* **208**: 301-308.

Farrand, S.K., Hwang, I., and Cook, D.M. (1996) The tra region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J Bacteriol* **178**: 4233-4247.

Fuqua, W.C., and Winans, S.C. (1994) A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J Bacteriol* **176**: 2796-2806.

Gonzalez, J.E., and Marketon, M.M. (2003) Quorum sensing in nitrogen-fixing rhizobia. *Microbiol Mol Biol Rev* **67**: 574-592.

Haase, J., Lurz, R., Grahn, A.M., Bamford, D.H., and Lanka, E. (1995) Bacterial conjugation mediated by plasmid RP4: RSF1010 mobilization, donor-specific phage propagation, and pilus production require the same Tra2 core

components of a proposed DNA transport complex. *J Bacteriol* **177**: 4779-4791.

Haase, J., Kalkum, M., and Lanka, E. (1996) TrbK, a small cytoplasmic membrane lipoprotein, functions in entry exclusion of the IncP alpha plasmid RP4. *J Bacteriol* **178**: 6720-6729.

Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol Biol* **25**: 989-994.

Hwang, I., Li, P.L., Zhang, L., Piper, K.R., Cook, D.M., Tate, M.E., and Farrand, S.K. (1994) TraI, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid N-acylhomoserine lactone autoinducer. *Proc Natl Acad Sci U S A* **91**: 4639-4643.

Kalogeraki, V.S., and Winans, S.C. (1997) Suicide plasmids containing promoterless reporter genes can simultaneously disrupt and create fusions to target genes of diverse bacteria. *Gene* **188**: 69-75.

Kaniga, K., Delor, I., and Cornelis, G.R. (1991) A wide-host-range suicide vector for improving reverse genetics in gram-negative bacteria: inactivation of the blaA gene of *Yersinia enterocolitica*. *Gene* **109**: 137-141.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., 2nd, and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175-176.

Lessl, M., Krishnapillai, V., and Schilf, W. (1991) Identification and characterization of two entry exclusion genes of the promiscuous IncP plasmid R18. *Mol Gen Genet* **227**: 120-126.

Li, P.L., Hwang, I., Miyagi, H., True, H., and Farrand, S.K. (1999) Essential components of the Ti plasmid trb system, a type IV macromolecular transporter. *J Bacteriol* **181**: 5033-5041.

Li, P.L., and Farrand, S.K. (2000) The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. *J Bacteriol* **182**: 179-188.

Lyras, D., Chan, A.W., McFarlane, J., and Stanisich, V.A. (1994) The surface exclusion system of RP1: investigation of the roles of *trbJ* and *trbK* in the surface exclusion, transfer, and slow-growth phenotypes. *Plasmid* **32**: 254-261.

Pappas, K.M., and Winans, S.C. (2003) A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol Microbiol* **48**: 1059-1073.

Peng, W.T., Banta, L.M., Charles, T.C., and Nester, E.W. (2001) The *chvH* locus of *Agrobacterium* encodes a homologue of an elongation factor involved in protein synthesis. *J Bacteriol* **183**: 36-45.

Piper, K.R., Beck von Bodman, S., and Farrand, S.K. (1993) Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. *Nature* **362**: 448-450.

Press, R., Glansdorff, N., Miner, P., De Vries, J., Kadner, R., and Maas, W.K. (1971) Isolation of transducing particles of phi-80 bacteriophage that carry different regions of the *Escherichia coli* genome. *Proc Natl Acad Sci U S A* **68**: 795-798.

Sambrook, J., and Russel, D.W. (2001) *Molecular Cloning - A laboratory manual.* New York: Cold Spring Harbor Laboratory Press.

Sukupolvi, S., and O'Connor, C.D. (1990) TraT lipoprotein, a plasmid-specified mediator of interactions between gram-negative bacteria and their environment. *Microbiol Rev* **54**: 331-341.

Tempe, J., Petit, A., Holsters, M., Montagu, M., and Schell, J. (1977) Thermosensitive step associated with transfer of the Ti plasmid during conjugation: Possible relation to transformation in crown gall. *Proc Natl Acad Sci U S A* **74**: 2848-2849.

Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J., Farrand, S.K., and Winans, S.C. (2000) The bases of crown gall tumorigenesis. *J Bacteriol* **182**: 3885-3895.

Zhu, J., and Winans, S.C. (2001) The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. *Proc Natl Acad Sci U S A* **98**: 1507-1512.

CHAPTER FOUR

RepC Protein of the Octopine-type Ti Plasmid Binds to the Putative Origin of Replication Within *repC* with High Specificity and Cooperativity[4]

4.1. Summary

Vegetative replication and partitioning of many plasmids and some chromosomes of alpha-proteobacteria are directed by their *repABC* operons. RepA and RepB ensure accurate partitioning of newly synthesized replicons to daughter cells, while RepC is the sole protein that is essential for replication initiation. RepC proteins do not resemble any characterized replication initiation protein. Here we show that the origin of replication resides within the *repC* gene by demonstrating that a plasmid containing just *repC*, an antibiotic resistance gene, and a ColE1 origin is able to replicate in *Agrobacterium tumefaciens*. Purified RepC protein bound to a site within the *repC* gene with moderate affinity, high specificity and cooperativity (Hill coefficient of 2). The binding site was localized by DNA resection and by DNase I protection assays to an AT-rich region that contains a large number of GANTC sites, whose methylation state may help to regulate replication. A fragment of RepC containing residues 26-158 is sufficient to bind this site, though affinity and specificity is somewhat lower than for the full length protein. This portion of RepC is predicted to have structural homology to members of the MarR family of transcriptional factors. Despite exhaustive efforts, we were unable to provide RepC *in trans* to a plasmid lacking this gene but containing *oriV*, suggesting that RepC may function only *in cis*. This was confirmed by showing that overexpression of RepC in *A. tumefaciens* caused

[4] Manuscript to be submitted for publication in the journal *Molecular Microbiology*.

large increases in copy number *in cis* but no change in copy number of other plasmids containing the same *oriV* sequence *in trans*.

4.2. Introduction

Plasmids are semiautonomous genetic elements that reside within many genera of bacteria, and can be found in all three domains of life (del Solar *et al.*, 1998). They are by definition non-essential for cell viability. Many confer no obvious benefit to their hosts and can be thought of as genetic parasites, while other plasmids provide their hosts with remarkable new survival strategies, including the abilities to detoxify various antimicrobials and other toxic substances (Nikaido, 2009), to utilize novel compounds as nutrients (Phale *et al.*, 2007), or to interact pathogenically or symbiotically with host plants or animals (Crossman, 2005; Johnson and Nolan, 2009). Plasmids have evolved a variety of mechanisms to ensure replication once per generation within their bacterial hosts and to accurately distribute newly replicated plasmids to daughter cells prior to cell division (Bingle and Thomas, 2001; Funnell, 2005; Thomas, 2000).

Most plasmids have an essential replication initiator protein that plays both positive and negative roles in replication initiation (Chattoraj, 2000; Paulsson and Chattoraj, 2006). These proteins bind to the origin of replication (*oriV*), which in many cases contains directly repeated DNA sequences called iterons (del Solar *et al.*, 1998). One of the ways that replication is limited is by restricting the concentration of the replicator protein, either via negative autoregulation, antisense RNA, protein instability, covalent modification, or dimerization (Paulsson and Chattoraj, 2006). In several plasmids, initiator proteins can form so-called “handcuffs” in which a dimer of the protein binds to the iterons of two sister plasmids, preventing further rounds of replication until the plasmids are pulled apart by the partitioning machinery (Das and

Chattoraj, 2004; Zzaman and Bastia, 2005). In some cases, initiators at low concentrations form active monomers, while at high concentration they form inactive dimers (Chattoraj, 2000).

The alpha-proteobacteria is a fascinatingly diverse group that includes photoautotrophs, saprophytes, mutualists, and pathogens. Plasmids play central roles in many if not most of these survival strategies. A prime example is found in the plant pathogen *Agrobacterium tumefaciens,* which causes crown gall tumors on dicotyledonous plants (Smith and Townsend, 1907). Pathogenesis requires a 200 kb plasmid called the Ti (tumor-inducing) plasmid, a portion of which is transferred into plant nuclei via a mechanism that clearly evolved from conjugative transfer (Alvarez-Martinez and Christie, 2009). This transferred DNA is integrated into the host genome and directs the production of phytohormones, causing plant cell proliferation, and bacterial nutrients called opines (Dessaux *et al.*, 1998). Opines are utilized by the bacteria via dedicated permeases and catabolic enzymes, all encoded by the plasmid (Zhu *et al.*, 2000). Ti plasmids can replicate at low or moderate copy number and are partitioned to daughter cells with extremely high fidelity (Pappas, 2008).

All Ti plasmids as well as many if not all other *Agrobacterium* plasmids and secondary chromosomes contain *repABC* operons (Slater *et al.*, 2009; Wood *et al.*, 2001). RepA and RepB proteins resemble components of plasmid partitioning systems that are widespread in bacterial plasmids, some lysogenic plasmids, and some chromosomes (Funnell, 2005). Both proteins are dispensable for plasmid propagation, though null mutations in either gene reduce the efficiency of plasmid inheritance, resulting in the accumulation of cured cells within a culture (Chai and Winans, 2005a). RepC is the only protein that is essential for plasmid propagation. A plasmid containing *repC* and a small amount of upstream DNA is capable of autonomous replication (Chai and Winans, 2005b; Izquierdo *et al.*, 2005). The origin of replication

must therefore lie within this region. To our knowledge no RepC protein from any bacterium has been studied biochemically.

Expression of the *repABC* operon of the octopine-type Ti plasmid is regulated in complex ways. First, there are four promoters upstream of *repA*. The P4 promoter provides basal levels of expression sufficient to ensure plasmid replication and partitioning (Pappas and Winans, 2003b). This promoter is also inducible by phosphorylated VirG, and is therefore stimulated by phenolic compounds that induce the *vir* regulon (Cho and Winans, 2005). Consequently phenolic compounds increase Ti plasmid copy number 5-10 fold. The other three promoters are activated by the quorum-sensing protein TraR (Pappas and Winans, 2003a). The quorum-sensing system composed of TraR and the pheromone synthase protein TraI can also elevate copy number. The operon is also negatively autoregulated. RepA binds directly downstream of P4 and represses transcription (Pappas and Winans, 2003b). RepB potentiates the ability of RepA to bind to this site by forming a RepA-RepB complex (Chai and Winans, 2005a). Conversely, RepB can bind to a site between *repA* and *repB* inhibiting transcription of *repB* and *repC*. RepA potentiates the ability of RepB to do so. RepA/RepB complexes binding to both sites may cause a protein-DNA loop that strongly represses the *repABC* operon. Finally, an antisense RNA, *repE*, is found directly upstream of *repC*. This small RNA hybridizes with complementary mRNA, blocking transcription and translation of *repC* (Cervantes-Rivera *et al.*, 2010; Chai and Winans, 2005b). In previous studies, we have shown that a plasmid containing a fragment with *repE* and *repC* can replicate in *A. tumefaciens* (Chai and Winans, 2005b). The plasmid containing just *repC* failed to do so, indicating either that (i) such a plasmid replicated at lethally high levels preventing colony formation, or (ii) *repE* plays some essential role in replication, or (iii) *oriV* lies within *repE*. RepE RNA

can also act *in trans* to downregulate other copies of *repC* in the same cell, and is therefore the major mediator of plasmid incompatibility (Chai and Winans, 2005b).

Another member of the alpha-proteobacteria whose cell cycle has received considerable attention is *Caulobacter crescentus*. This organism has a DNA methylase, CcrM, that methylates the adenine residue of GANTC sites (Collier *et al.*, 2007). The abundance of this methylase is cell-cycle regulated, with a burst of synthesis and activity directly before replication initiation. This causes hemimethylated sites to become fully methylated, which appears to alter the activity of several promoters involved in the cell cycle (McAdams and Shapiro, 2003). *A. tumefaciens* has a gene predicted to encode an orthologous methylase (Slater *et al.*, 2009; Wood *et al.*, 2001). It may be significant that there are five tightly clustered GANTC sites near the middle of the *repC* gene, and three more in the promoter region of *repE* (Fig. 4.1). This is reminiscent of the chromosome origin of *E. coli* which is rich in GATC sites, which are methylated by Dam. Methylation of GATC sites is delayed due to binding of the SeqA protein, which is thought to sequester these sites from Dam (Waldminghaus and Skarstad, 2009).

As described above, virtually nothing is known about the biochemical properties of any RepC protein. The origin of replication has not been identified, although it must lie within *repC* or directly upstream. There are no directly repeated DNA sequences anywhere in this region, so replication evidently does not require iterons. However, there is an AT-rich region within *repC* that could be part of a replication origin. In an effort to understand this protein, we purified it from an overexpressing strain of *E. coli*, and tested it for binding to sites within *repE* and *repC*. We have also tested fragments of RepC protein for DNA binding. Finally, we tested the ability of overexpressed RepC to function *in cis* and *in trans*. To our knowledge

this is the first study showing interaction of a purified RepC protein with the putative origin of replication among the large group of *repABC* replicons.

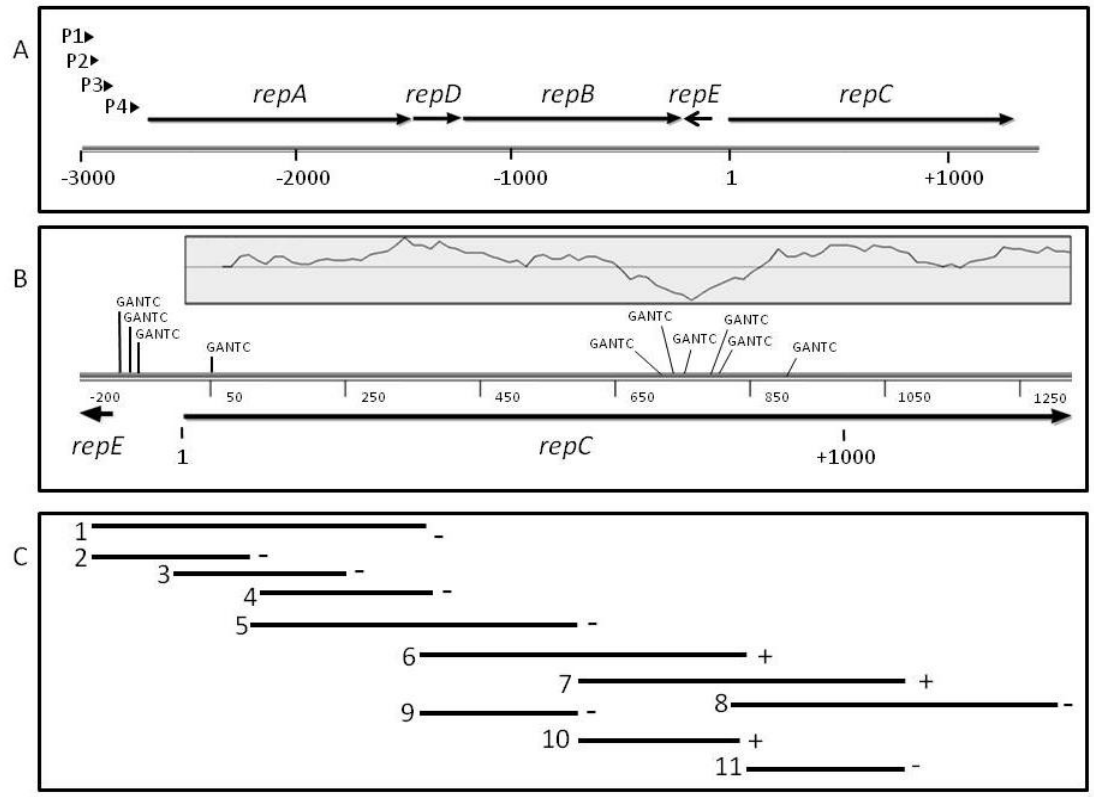


Figure 4.1. Genetic map of the *repABC* region.
A. RepA and RepB resemble plasmid partitioning systems. RepC is the only essential replication initiator. RepE is an antisense RNA that downregulates RepC expression. RepD has no known function. Promoters P1, P2, and P3 are activated by TraR-OOHL complexes, while P4 provides basal levels of transcription, is activated by phospho-VirG, and is autorepressed by RepA and RepB.
B. Genetic map of *repC* showing the clusters of GANTC sites at the *repE* promoter and at the middle of *repC*. The GC content is shown in the box above *repC*, and shows an AT-rich region near the middle of the gene.
C. DNA fragments tested for binding of RepC. "+" and "-" to the right of each fragment indicate whether that fragment was bound by RepC.

4.3. Results

oriV lies within *repC*

Plasmids lacking *repE* failed to transform *A. tumefaciens* though this could have been due to lethal levels of replication (Chai and Winans, 2005b; Izquierdo *et al.*, 2005). We therefore constructed a plasmid lacking all *repE* sequences and containing a P*lac*-*repC* fusion, and an antibiotic cassette. No extra sequence upstream or downstream of *repC* was inserted into this construct to make sure that *repC* alone contained all the necessary *cis* acting elements required for replication. The *lacI* gene was expressed on a separate plasmid. By adding IPTG, the P*lac* promoter was expressed at levels appropriate for non-lethal expression of *repC*, and the plasmid replicated in *A. tumefaciens* (Table 4.1), forming colonies under conditions that select for the plasmid, and containing readily detectable levels of plasmid DNA (result not shown). The origin of replication lies within the *repC* gene and does not depend on the presence of any extra-genic *cis*-acting elements. We also showed that an intact RepC protein is necessary, as truncations generated by frameshift mutations in *repC* rendered the plasmid unable to replicate (Table 4.1).

Table 4.1. Electroporation of UIA143(pSRKKm) with the indicated plasmids.

		[IPTG] (mM)				
Plasmid	Description	0.0	0.1	1	5	25
pUP001	Vector (ColE1 *ori*, Sp^R *lacZ*α)	-	-	-	-	-
pUP003	pUP001 + *repC*	-	-	Colonies	Colonies	Colonies
pUP004	pUP003 lacking P*lac* promoter	-	-	-	-	-
pUP005	pUP003 with frameshift mutation in codon 303	-	-	-	-	-

Purification of RepC protein

Our first attempts to purify soluble RepC either on its native form or as his-tagged fusions were not successful. However, a His_6-MPB-RepC fusion protein provided a good yield (1 mg/ml) of highly soluble protein. In early experiments, the fusion protein was used intact. In later experiments, the His_6-MBP portion of the fusion was removed by cleavage with TEV protease followed by IMAC chromatography (see Experimental procedures), yielding highly pure RepC protein that has a wild type sequence except for three amino acids at the N-terminus. We got similar results with both protein preparations.

Localization of a RepC binding site within the *repC* gene

oriV most likely lies within the *repC* or *repE* genes (Bartosik *et al.*, 1998; Chai and Winans, 2005b; Izquierdo *et al.*, 2005; MacLellan *et al.*, 2005; Venkova-Canova *et al.*, 2004). We used electrophoretic mobility shift assays (EMSA) to test six overlapping fragments of the *repC* region for the ability to be bound by MBP-RepC. Each fragment was 500 nucleotides in length (Fig. 4.1). All fragments seemed to be shifted weakly at high protein concentrations (results not shown). However, fragments 6 and 7 were shifted completely at lower RepC concentrations. Fragment 1 seemed to have some affinity, so we subdivided it into three overlapping fragments, and found that none of them was shifted (Fig. 4.1). We concluded that the long length of the six original fragments tended to give spurious results and that there was a bona fide binding site in the middle of the gene.

Fragments 6 and 7 cover 750 nucleotides and share a region of 250 nucleotides (Fig. 4.1). We PCR amplified three 250 nucleotide fragments covering this region (fragments 9, 10, and 11), and found that only fragment 10 was bound by MBP-RepC.

We then narrowed the binding site by synthesizing various PCR fragments within this 250 nucleotide region (Fig. 4.2).

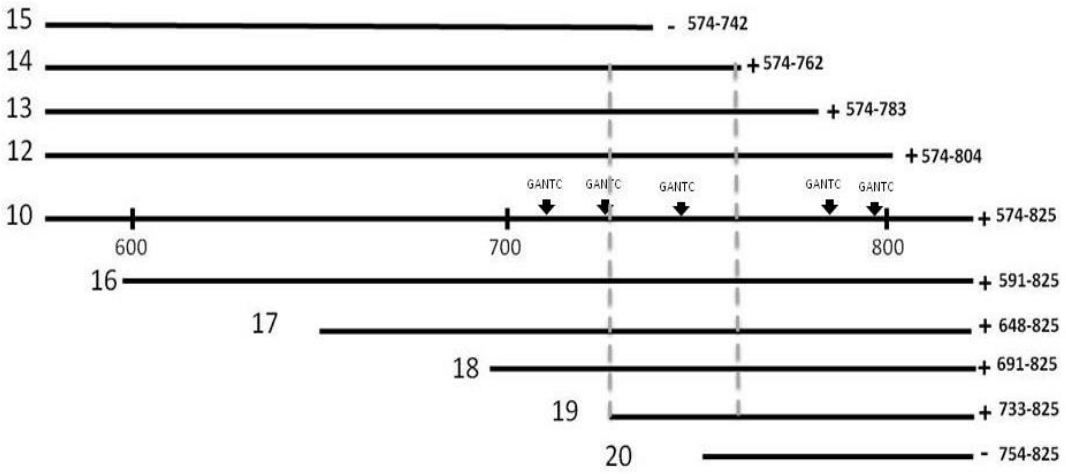


Figure 4.2. Mapping the RepC binding site starting with fragment #10 from Figure 4.1. Note the position of the GANTC sites around the binding site. "+" and "-" to the right of each fragment indicate whether that fragment was bound by RepC. Each fragment is marked with the relative position from nucleotide 1 of *repC* gene.

The binding site was further localized by using DNase I footprinting. Fragment 10 (Fig. 4.1) was PCR amplified with a fluorophore at one 5' end. To this PCR product, we added MBP-RepC (or RepC) and DNaseI, and submitted the resulting fragments for automated mapping of the DNase I cleavage sites (see section 4.5 Experimental Procedures). Either MBP-RepC (Fig. 4.3) or just RepC (result not shown) protected a region of approximately 39 nucleotides positioned exactly where the binding site was predicted to lie as judged by the EMSA studies. An imperfect dyad symmetry is found within this region. It is possible that RepC binds to this site as a rotationally symmetric multimer.

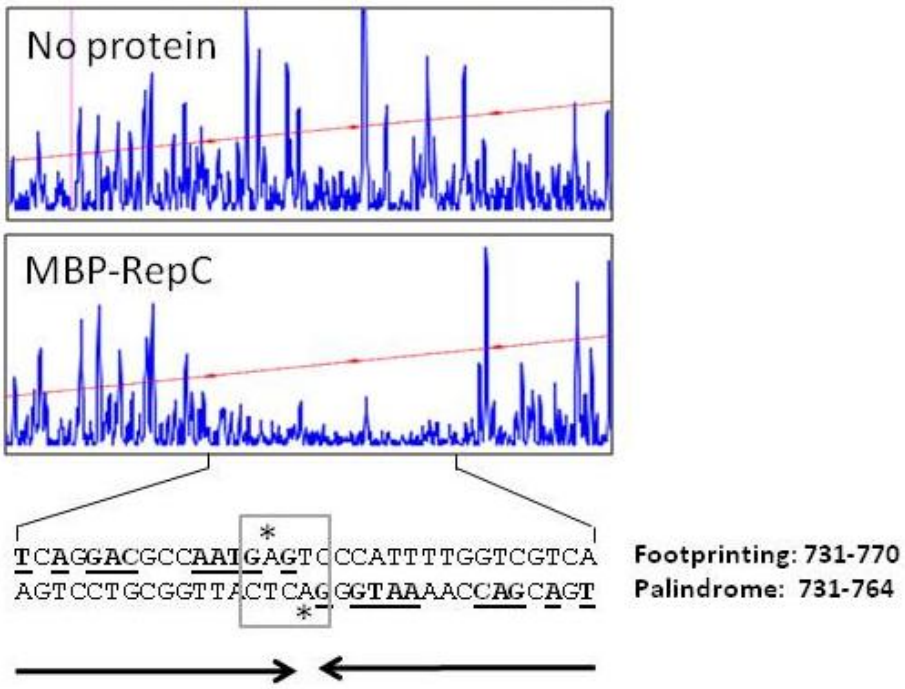


Figure 4.3. DNase I footprinting of the RepC binding site. The protected sequence is shown at the bottom. Dyad symmetrical DNA sequences are underlined, and a GANTC site is boxed.

DNA binding is cooperative and uninfluenced by methylation of GANTC sites

In studying MBP-RepC binding via EMSA, we noticed that binding appeared to be cooperative. Often, a fragment containing the binding site appeared to be completely bound by one protein concentration, but almost completely unbound when the protein was just 2-fold more dilute. We addressed this question using several representative fragments, and diluting MBP-RepC or RepC in 1.5 fold increments. The results confirmed that DNA binding is cooperative, with a Hill coefficient of approximately 2 (Fig. 4.4). It therefore seems plausible that RepC can multimerize upon binding to DNA. It is also possible that RepC-DNA complexes contain two DNA fragments rather than one and that this multimerization helps to drive the observed cooperativity.

As described above, *repC* has a cluster of GANTC sites near the center of the gene. It seemed plausible that methylation could affect protein affinity. We therefore cloned the gene that directs GANTC methylation in *A. tumefaciens*, designated *ccrM*, and expressed the methylase in strains of *E. coli* that also contained a plasmid with the RepC binding site (fragment 10, see Fig. 4.2). Methylation of GANTC sites was confirmed by using the restriction endonuclease HinfI, which digests unmethylated GANTC sites (results not shown). Methylated and unmethylated DNA fragments were subjected to binding assays with RepC, but no difference in binding affinity was observed between fully methylated and unmethylated DNA.

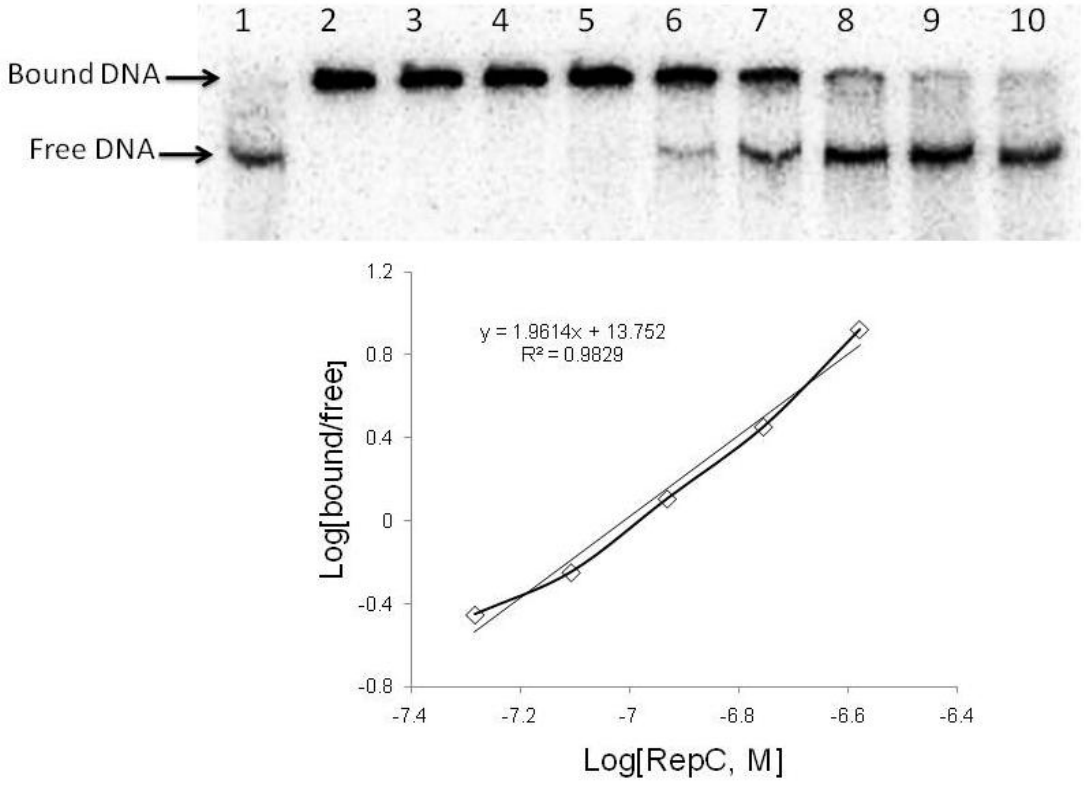


Figure 4.4. Electrophoretic mobility shift assay of RepC binding to a fragment containing the binding site. The slope of the graph provides the Hill coefficient, a measurement of binding cooperativity. A Hill coefficient greater than 1 indicates positive cooperativity. RepC was provided at 1 μM in lane 2, and was diluted 1.5 fold serially in each of the other lanes. Fragment 10 was used in this assay (Fig. 4.2). k_D for RepC biding to this fragment is 0.1 μM.

We reasoned that a comparison between hemi-methylated and fully methylated DNA would have more biological meaning. Therefore, we tested oligos 42 nucleotides in length that span the binding site and which contain N6-methyl-dA in place of dA (IBA GmbH, Germany). The two strands were hybridized with each other and with the identical sequence lacking methyl groups, and tested for RepC binding. We performed binding assays by using EMSA to determine the RepC binding affinity. However, we did not detect any difference in binding affinity (Fig. 4.5). We conclude that methylation of GANTC sites does not alter binding affinity of purified RepC.

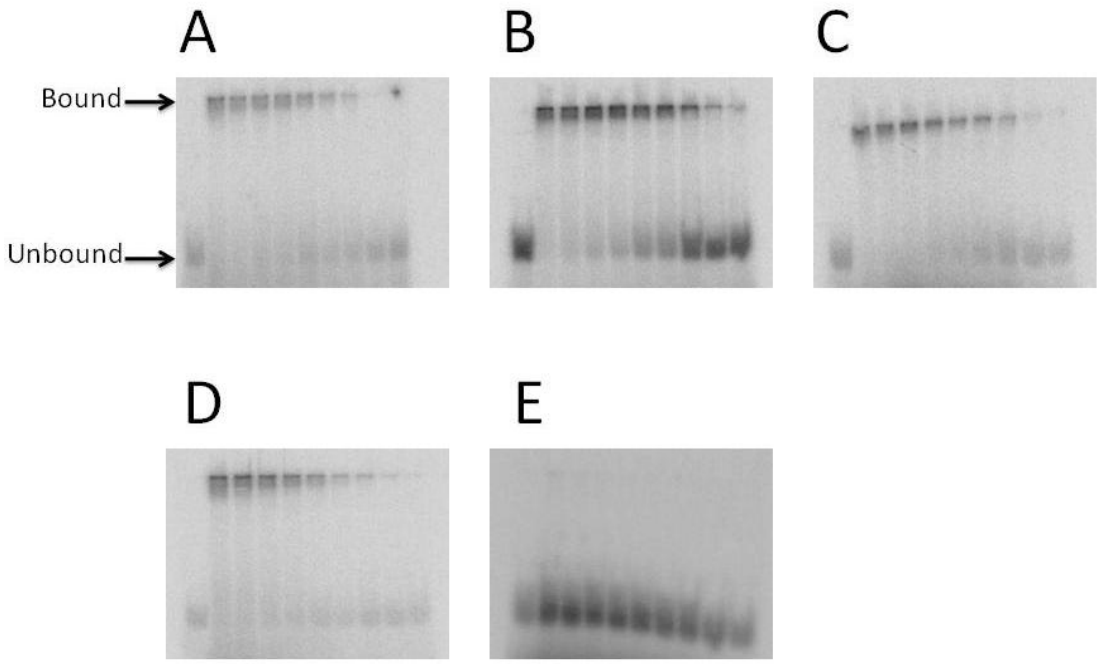


Figure 4.5. Electrophoretic mobility shift assays for RepC binding onto 42 bp DNA fragments. k_D for binding was 0.1 μM for gels A-D. A – non methylated DNA substrate; B – bottom strand methylated; C – top strand methylated; D – both strands methylated; E – *oriV* mutated sequence (Table 4.5).

We also tested whether RepC binding could be affected by the presence of ATP in the binding reactions. ATP enhances binding of RepA to its binding site (Pappas and Winans, 2003b) and it also has a crucial role in DnaA binding at *oriC*

(Ozaki and Katayama, 2009). We did not detect any change in RepC binding affinity when ATP was added to the binding reactions (results not shown).

The N-terminal half of RepC is sufficient for high affinity DNA binding

In an effort to identify structural homologs of RepC, we used the Phyre web server to compare its sequence to those of all proteins whose structures have been solved (Kelley and Sternberg, 2009). This algorithm detected a significant structural similarity between RepC and several proteins with a winged helix-turn-helix DNA binding domain such as the OhrR protein of *B. subtilis*. OhrR has been visualized as a complex with its DNA binding site (Hong *et al.*, 2005). The similarity was limited to residues 38-186 of RepC (Fig. 4.6). Homologous residues of OhrR include its DNA recognition helix and dimerization determinants.

The homology between OhrR and the N-terminal half of RepC led us to wonder whether this half of RepC might suffice for DNA binding. To test this, we fused a number of *repC* fragments to *malE* and purified the resulting fusion proteins. Many of them did bind to the same region of DNA as the full length protein. A fragment containing resides 1-158 of RepC was proficient in binding, while a fragment containing residues 1-155 was deficient. Likewise, a fragment containing residues 26-170 was binding proficient, while a fragment containing residues 30-170 was deficient (Fig. 4.7). Affinities of these fragments were in general lower than those of full length RepC. A fragment containing the C-terminal domain of RepC was defective for DNA binding (data not shown).

Several truncated RepC proteins caused the protein to poorly discriminate between specific and non-specific sequences. Proteins containing RepC residues 1-320, 1-200, and 1-190 retained the ability to discriminate against specific and non-specific DNA, while proteins containing RepC residues 1-158, 1-161, 1-171, 1-175,

and 180 bound non-specific DNA with k_Ds similar to the ones for specific DNA (Fig. 4.7 and data not shown). The reasons are unclear.

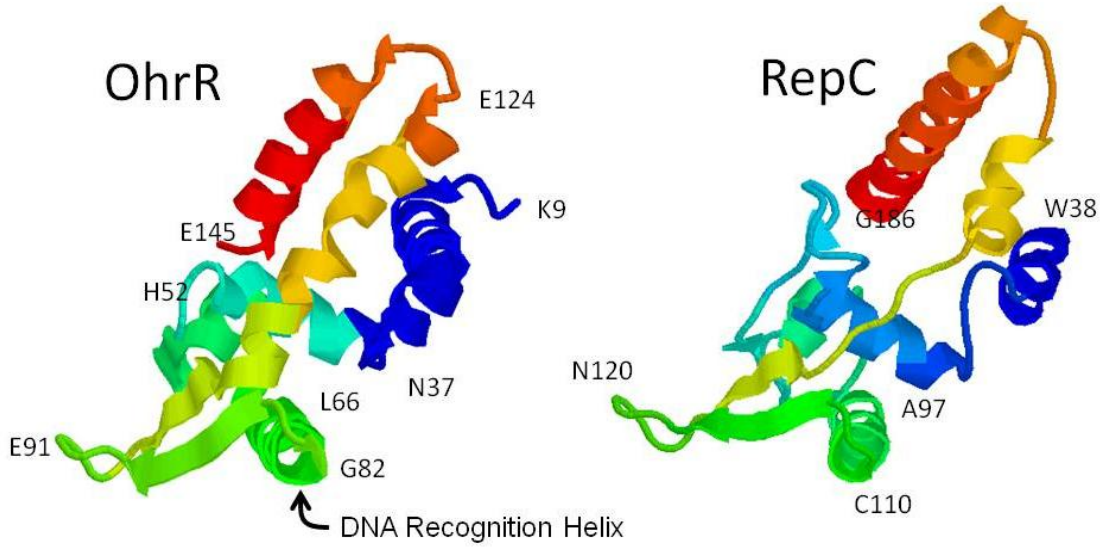


Figure 4.6. RepC presents structural homology to the MarR transcriptional regulators (OhrR shown in this case), as determined by using the Phyre protein fold recognition server (http://www.sbg.bio.ic.ac.uk/~phyre/).

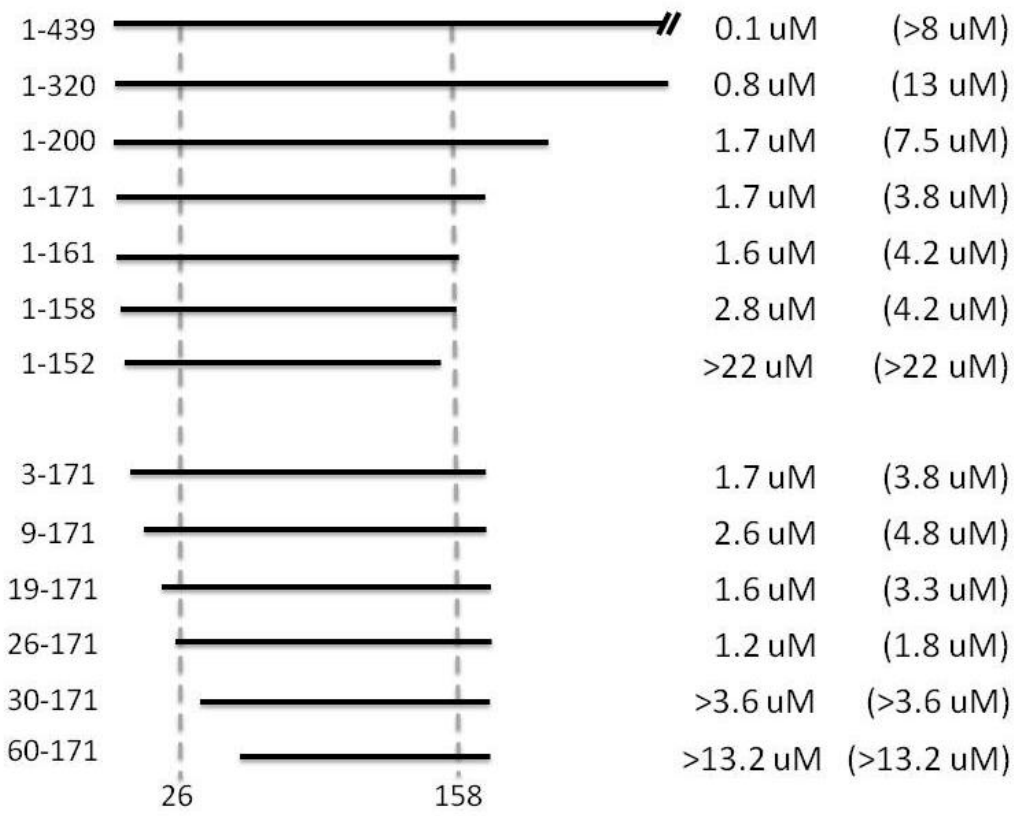


Figure 4.7. Fragments of RepC protein tested for binding. k_Ds are provided for the specific DNA fragment 10 (Fig. 4.2). k_Ds for non-specific DNA sequences are shown in parentheses. DNA binding domain was located in the NTD of RepC (dashed vertical lines). The sign ">" indicates a k_D higher than the value shown since no DNA binding was detected at that protein concentration.

Overexpression of RepC causes strong overreplication of RepC-dependent plasmids

We have made conprehensive efforts to identify a functional *oriV*. Essentially we subcloned fragments of the *repE-repC* region into narrow host range plasmids, and attempted to introduce the resulting plasmids into a strain of *A. tumefaciens* in which RepC was expressed at various levels *in trans*, and selecting for plasmid replication in *A. tumefaciens*. In some constructs, we attempted to ensure that the *oriV* was transcribed, as the origin of some plasmids functions only if transcribed (Bruand and Ehrlich, 1998; del Solar *et al.*, 1998). These efforts have so far not succeeded (Table 4.2). We have also taken advantage of plasmid pYC212 which has a mutation in the *repE* promoter that decreases RepE expression, and increases *repC* expression, thereby elevating the plasmid copy number about 20 fold (Chai and Winans, 2005b). Strain UIA143(pYC212) was electroporated with suicide vectors containing different fragments of *repC* (some containing the entire sequence of the gene), but none of them was able to replicate (as judged by the absence of colonies on selective media, results not shown). These results lead us to speculate that perhaps RepC functions preferentially or solely *in cis*. It may be noteworthy that RepA, RepB, and TraA all function preferentially *in cis* (Cho and Winans, 2005; Pappas and Winans, 2003b).

Table 4.2. Electroporation of UIA143(pUP500)[1] with the designated plasmids[5].

		[IPTG] (mM)				
Plasmid	Description	0.0	0.1	1	5.0[2]	25[2]
pUP001	Vector (ColE1 *ori*, Sp^R *lacZ*α)	-	-	-	-	-
pUP003	pUP001 + *repC*	-	-	Colonies	-	-
pUP004	pUP003 lacking P*lac* promoter	-	-	-	-	-
pUP005	pUP003 with frameshift mutation in codon 303	-	-	-	-	-
pBluescript SK^+	Vector (ColE1 *ori*, Ap^R *lacZ*α)					
pUP502	pBluescript SK+ repC frameshift in codon 88.	-	-	-	-	-
pUP509	pBluescript SK+ repC frameshift in codon 303.	-	-	-	-	-

[1]pUP500 is a derivative of pSRKKm expressing a P*lac-repC* fusion, which is regulated by the LacI protein encoded by the vector. [2] The absence of colonies for these plates indicates that an excess of RepC expression may have led to runaway replication of plasmid pUP500.

In an effort to determine whether RepC might function only *in cis*, we took advantage of the fact that overexpression of RepC causes increased plasmid copy number (Chai and Winans, 2005b; Cho and Winans, 2005; Li and Farrand, 2000; Pappas and Winans, 2003a). Mild overexpression of the native *repABC* operon by TraR or VirG causes a mild increase in copy number, while an unregulated P*lac-repC* fusion leads to runaway replication, which is lethal to the host (Chai and Winans, 2005b; Izquierdo *et al.*, 2005). We constructed two new P*lac-repC* plasmids in the broad host range plasmid pPZP201. The first of these, pUP450, retains the weak native *repC* ribosome binding site (RBS). In this plasmid, translation may be further weakened by translational occlusion, as the *lacZ* translation start site would cause ribosomes to cross the native *repC* RBS, translating in a different reading frame. The second construct, pUP455, contains a translational fusion between amino acids 1-14 of

[5] In a similar experiment, we used strain UIA143(pUP450), which shows an increase in plasmid copy number *in cis* (Figure 4.9), in the hope that the extra RepC expressed from this vector could rescue replication of constructs containing frameshift mutations on *repC* of plasmids pUP502 and 509. However we got no colonies. In a control experiment with pYC183 we got hundreds of colonies using similar plasmid concentrations.

lacZ and *repC*. RepC translation in this plasmid begins at the strong *lacZ* ribosome binding site. In initial experiments, these were introduced into an *A. tumefaciens* strain lacking the Lac repressor. Electroporation of the vector control (pPZP201) gave rise to numerous colonies, as expected. Electroporation of pUP450 also yielded numerous colonies that grew at slightly slower rates. In contrast, electroporation of pUP455 yielded no colonies, despite it having a separate wide host range replication system. In a control experiment, all three plasmids gave rise to colonies when transformed into *E. coli*, and plasmid concentration measurements indicated that DNA amounts were comparable. The most likely explanation is that pUP455 expresses high amounts of RepC, leading to lethal runaway replication, while pUP450, with its weaker RBS, made sub-lethal amounts of RepC.

To test whether pUP455 was expressing lethal amounts of RepC protein, we controlled *repC* expression by providing the Lac repressor *in trans*, using pSRKKm (Khan *et al.*, 2008). In the presence of plasmid pSRKKm, cells containing pUP455 formed colonies that grew at normal rates (Table 4.3). However, in the presence of high levels of IPTG (high levels are known to be required to inactivate Lac repressor in *A. tumefaciens* (Chen and Winans, 1991), cells containing pUP455 formed slow-growing colonies, while cells containing pUP450 formed normal growing colonies (Table 4.3). Colonies containing pUP455 and pSRKKm were inoculated into broth containing or lacking 5 mM IPTG. Addition of IPTG caused a strong increase in copy number of pUP455 (Fig. 4.8). Taken together, these findings confirm that overexpression of *repC* causes such high levels of replication that host viability can be impaired.

Table 4.3. Controlled expression of RepC prevents lethal runaway replication.

Source of Lac Repressor	pPZP201 derivative	Plates without IPTG	Plates containing 5 mM IPTG	Change in plasmid concentration with IPTG
none	pPZP201 (no *repC*)	Colonies	NA	NA
none	pUP450 (*repC* expressed weakly)	Colonies	NA	NA
none	pUP455 (*repC* expressed strongly)	No colonies	NA	NA
pSRKKm	pPZP201 (no *repC*)	Colonies	Colonies	No change
pSRKKm	pUP450 (*repC* expressed weakly)	Colonies	Colonies	No change
pSRKKm	pUP455 (*repC* expressed strongly)	Colonies	Small colonies	Strong increase

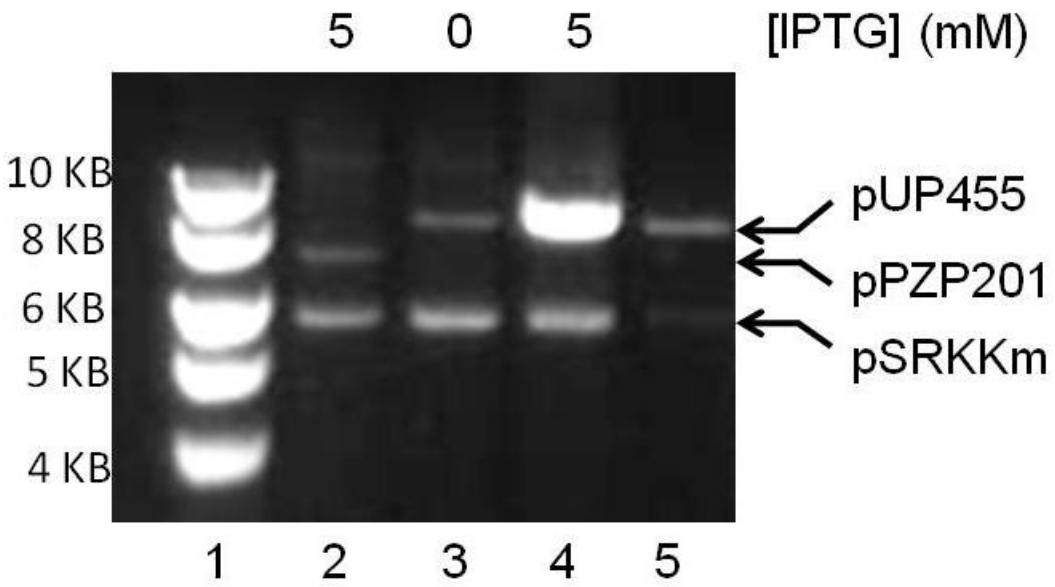


Figure 4.8. Addition of IPTG to strains expressing Lac repressor increases the copy number of pUP455.
Gel of digested plasmid DNA with *Kpn*I which generates a band of 7.1 Kb for pPZP201, 8.5 kb for pUP455 and 5.8 kb for pSRKKm. All strains contain pSRKKm, which expresses Lac repressor and serves as an internal control for plasmid extraction. The gel shows equal amounts of pSRKKm, and an increased amount of pUP455 upon IPTG addition. Lane 1: molecular mass standards; Lane 2: pPZP201 + 5 mM IPTG; Lane 3: pUP455; Lane 4: pUP455 + 5 mM IPTG; Lane 5: 1/10 dilution of lane 4.

RepC functions only *in cis*

Armed with the knowledge that overexpression of RepC causes severe overreplication of RepC-dependent plasmids, we asked whether overexpression of RepC on one plasmid would affect the copy number of a separate RepC-dependent plasmid. This was done in three ways. First, pUP450 was introduced into a strain carrying pYC183 which has a Plac-*repC* fusion, but which replicates at low levels, due to the presence of *repE* which downregulates expression of RepC (Chai and Winans, 2005b). The copy number of pYC183 was the same in a strain containing pUP450 as in a strain carrying a vector control (Fig. 4.9.A and Fig. 4.10). Second, similar results were obtaining using pKP23 in place of pYC183 (Fig. 4.9.B). Plasmid pKP23 contains the entire *rep* operon. Third, we did similar experiments using pUP455 in place of pUP450. Plasmid pUP455 is lethal unless Lac repressor is provided (on pSRKKm), and its copy number is increased by IPTG (Fig. 4.8). Addition of IPTG elevated the copy number of pUP455 (as seen above) but had no effect on the copy number of pKP23 (Fig. 4.9.C). We conclude that RepC is *cis*-acting, at least when overexpressed. This *cis*-activity could help to explain why we could not identify a functional *oriV* fragment *in trans*.

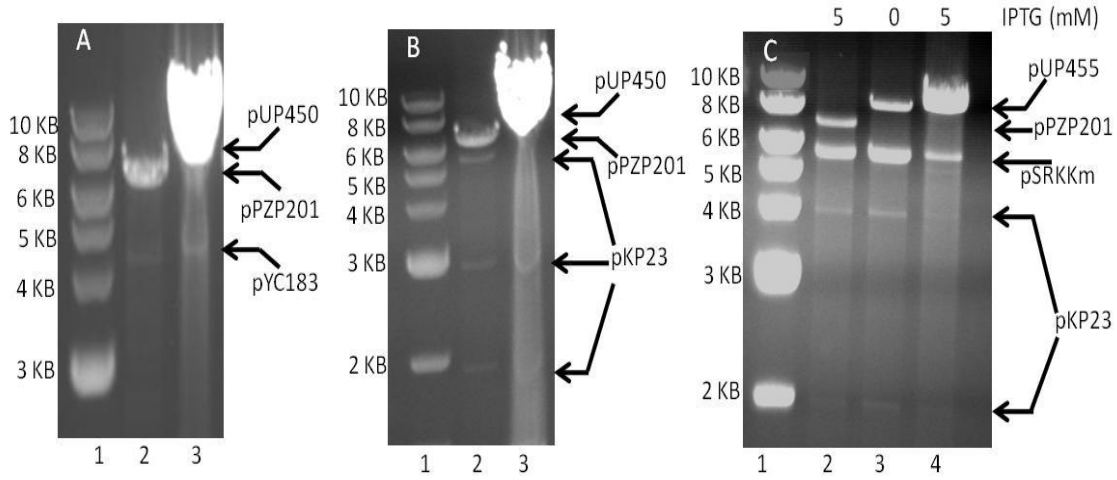


Figure 4.9. High-level expression of RepC on one plasmid does not affect the copy number of a separate RepC-dependent plasmid.
A. Lane 1: molecular mass standards. Lane 2: pPZP201 and pYC183. Lane 3: pUP450 and pYC183. pUP450 expresses RepC at high but non-lethal levels, and replicates at higher levels than the parent plasmid pPZP201. However, a second plasmid, pYC183, which contains *repC* and is RepC-dependent for replication, is not affected by RepC overexpression by pUP450. *Kpn*I and *Hind*III digestion.
B. Lane 1: molecular mass standards; Lane 2: pPZP201 and pKP23; Lane 3: pUP450 and pKP23. pKP23 contains *repABC* and requires RepC for replication. High level expression of RepC in pUP450 elevates the copy number of pUP450 (*in cis*) but fails to alter the copy number of pKP23. *Kpn*I and *Hind*III digestion.
C. Lane 1: molecular mass standards; Lane 2: pPZP201 and pSRKKm and pKP23 in the presence of 5 mM IPTG; Lane 3: pUP455 and pSRKKm and pKP23 without IPTG; Lane 4: same as lane 3 but after growth in the presence of 5 mM IPTG. pSRKKM provides Lac repressor. pKP23 contains *repABC* and requires RepC for replication. IPTG induces expression of RepC on pUP455, increasing copy number, but copy number of pKP23 is unaffected. Digestion with *Hind*III.

We confirmed the importance of the RepC binding site by designing a mutation that altered the DNA sequence at 16 positions across the RepC binding site without changing the amino acid sequence of the protein. We performed gel shift assays with the mutated DNA sequence and confirmed that RepC no longer recognizes the altered origin (Fig. 4.5.E). The mutations were introduced into plasmid pYC183 (which contains a ColEI origin and the *repC-repE* genes), creating plasmid pUP183(oriM1). This plasmid was unable to replicate in strain UIA143, as expected. The same mutation was cloned into vectors pUP450 and pUP455 (which contain a broad host range origin and two different P*lac-repC* fusions), creating vectors pUP450(oriM1) and pUP455(oriM1). Unlike the parental plasmids pUP450 or pUP455, neither plasmid showed elevated plasmid copy number (Fig. 4.10). We reasoned that by mutating the binding site *in cis*, we could overexpress RepC more strongly, and that perhaps it might then recognize a wild type *oriV in trans*. In order to test this, we electroporated strain UIA143 with vectors pUP450(oriM1) and several other plasmids containing all or part of *repC*, all having the RepC binding site. Among them, pUP502 is similar to pYC183 except that it contains a frameshift mutation near the 5'-end of the *repC* gene, and plasmid pUP509, which is also like pYC183, but contains a frameshift mutation downstream of the RepC binding site, ensuring normal levels of transcription across *oriV*. None of these plasmids were able to replicate on their own, as judged by the absence of colonies on selective medium when electroporated into strain UIA143 (results not shown). These plasmids were electroporated into UIA143(pUP450) and UIA143(pUP450(oriM1)) strains. None of them gave rise to colonies on selective medium. These results indicate that even when RepC is overexpressed, it still cannot function *in trans*.

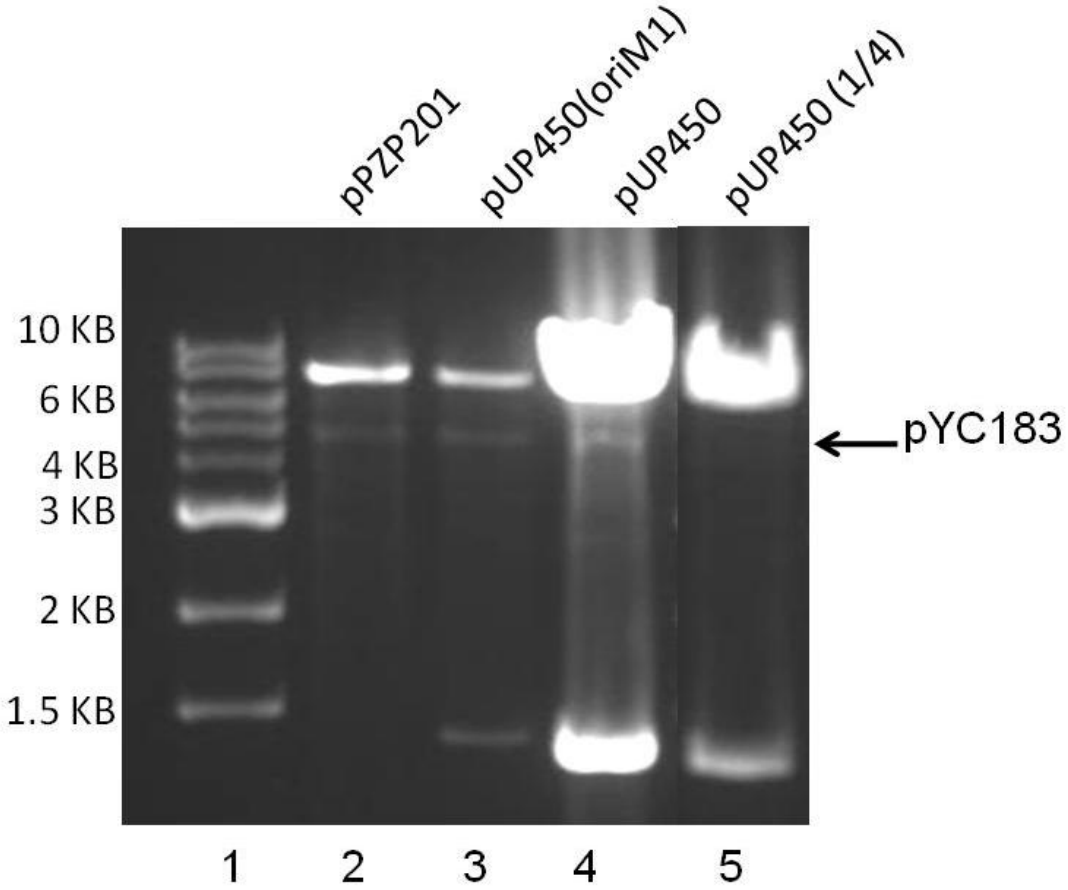


Figure 4.10. Mutation of the RepC binding site *in cis* does not improve RepC preference *in trans*.
Lane 1: molecular mass standards; Lane 2: pPZP201 and pYC183; Lane 3: pUP450(oriM1) and pYC183; Lane 4: pUP450 and pYC183; Lane 5: ¼ dilution of lane 4 to show that the fragment of pYC183 cannot be seen and therefore is unaffected by RepC overexpression from pUP450. pUP450(oriM1) expresses RepC that does not bind to its own origin. A second plasmid, pYC183, which contains *repC* and is RepC-dependent for replication, is not affected by RepC overexpression by either pUP450 or pUP450(oriM1). Arrow indicates restriction fragment of pYC183. *Kpn*I and *Hind*III digestion.

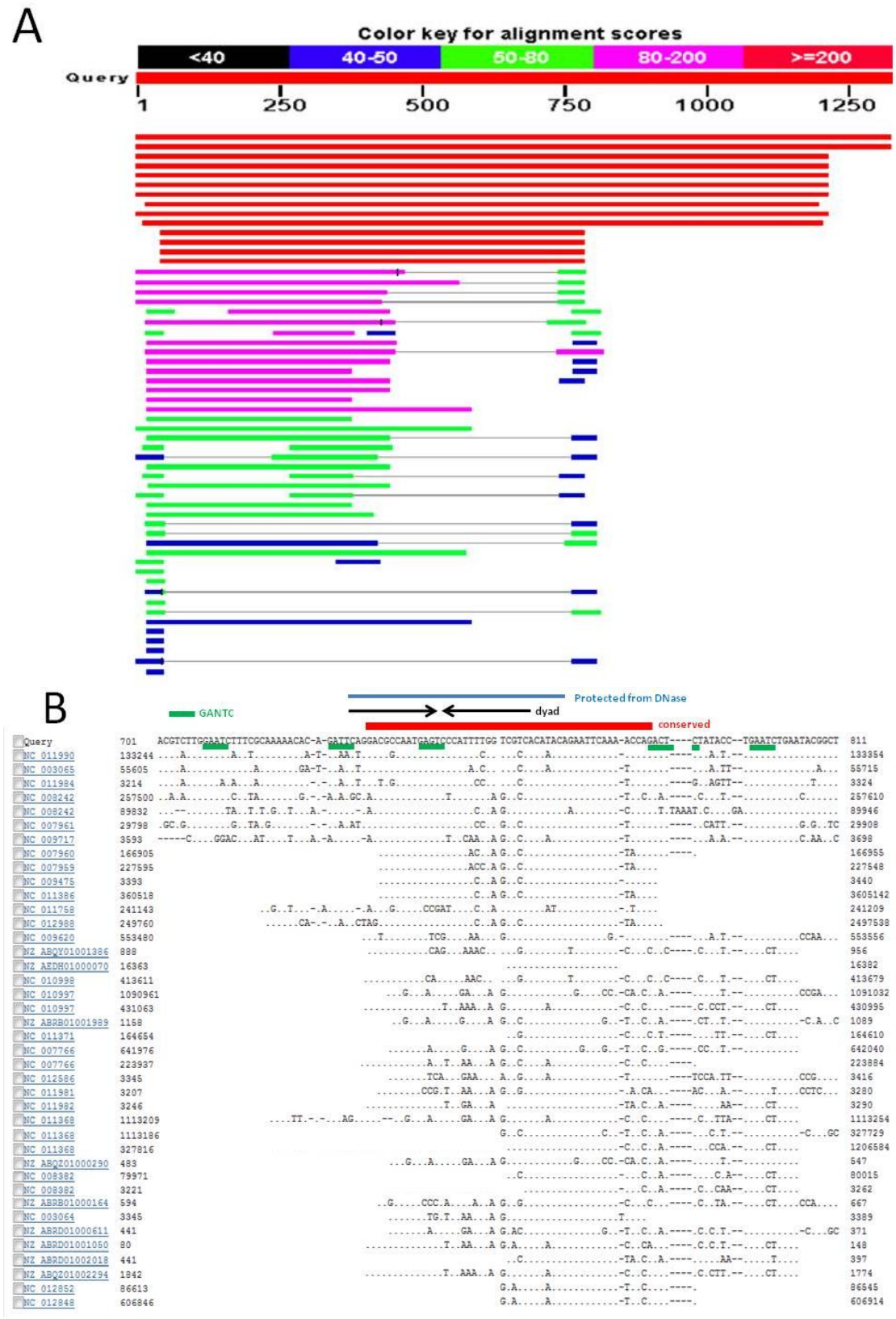
4.4. Discussion

One of the original goals of this study was to identify a DNA fragment containing all *cis*-acting sites essential for replication. We have succeeded in the sense that all of *repC* is required *in cis* for replication (Table 4.1). For a large number of plasmids, this is not the case, as it has been possible to provide a replication initiator *in trans* to the *oriV*. Ravin and collaborators (2003) reported that the replication protein RepA from plasmid prophage N15 acts only *in cis*, a property also observed in other replication proteins from phage P2 and ϕX-174. It may be relevant to mention that the RepA, RepB, TrbK/TrbJ and TraA proteins of the Ti plasmid also function preferentially *in cis* (Cho and Winans, 2007; Cho *et al.*, 2009; Pappas and Winans, 2003b). Conversely, the small RNA encoded by *repE* functions efficiently *in trans* to limit replication and mediate plasmid incompatibility (Chai and Winans, 2005b). *Cis* preference has been described for many proteins including transposases, in addition to partitioning and replication initiator proteins (Cho and Winans, 2007; Derbyshire and Grindley, 1996; McFall, 1986; Pappas and Winans, 2003b). The *cis*-preference might be due to low protein expression, proteolytic instability, poor diffusion within the cell, protein trapping by interactions to non-specific DNA, or to the cell membrane (Derbyshire and Grindley, 1996). A *cis* preference for RepC may also be due to the nascent RepC peptide getting trapped in its binding site during translation or immediately after it is released from the ribosome. In light of the conservation of the *repC* binding site (Fig. 4.11), it would make sense that the protein were to work preferentially *in cis* so to avoid binding to the wrong replication origin. Considering the multipartite genomes of many alpha-proteobacteria containing up to 7 *repABC* operons in a single cell (Castillo-Ramirez *et al.*, 2009; Cevallos *et al.*, 2008; Pappas, 2008) this makes even more sense, as a *cis*-acting replicator would improve chances of compatibility among so many replicons. *repABC* replicons of *A.*

tumefaciens tend to localize at the cell pole, but rarely do they co-localize (Kahng and Shapiro, 2003). One could envision that replicon localization together with *cis*-activity of partitioning and replication proteins would have an additive advantage on plasmid compatibility.

The fact that RepC appears to work only *in cis* makes it challenging to analyze the site upon which it acts. We are not certain that the RepC binding site comprises the entire *oriV* or part of it. However, in a review article (Cevallos *et al.*, 2008), another group working on a RepABC system in *Rhizobium etli*, reported physical mapping of the *oriV* of plasmid p42d to the middle of *repC*, although the data have not yet been published. As pointed out by those authors, many plasmids and bacteriophages have an *oriV* site embedded within the replication initiator gene. These include N15 prophage linear plasmid of *E. coli*, pAD1 of *Enterococcus faecalis*, pSX267 of *Sthaphylococcus xylosus*, pSK41 of *Staphylococcus aureus*, pLS32 of *Bacillus subtilis*, as well as bacteriophage lambda (Francia *et al.*, 2004; Gering *et al.*, 1996; Kwong *et al.*, 2004; Ravin *et al.*, 2003; Scherer, 1978; Tanaka *et al.*, 2005). Each of these origins contains short direct repeats called iterons that serve as binding sites for the replication initiator. In contrast, there are no iterons in the RepC binding site or anywhere else within or near *repC*.

Figure 4.11. Conservation of the *repC* DNA sequence.
A. Color illustration from a BLAST search showing conservation at the 5'-end and at *oriV* regions (around the middle of the gene) on *repC*. The 3' end of the gene shows DNA conservation only for a few members. Octopine-type DNA *repC* sequence was used for the BLAST search (blastn). Figure is not intended to provide details about specific genes.
B. Compilation of DNA sequences from the BLAST search showing conservation at and near the *repC* binding site. Note the strong conservation that starts in the DNaseI protected region and extends about 20 nucleotides downstream of it. Query sequence is shown in the first lane. Dots represent identical nucleotide sequences.


A
Color key for alignment scores
<40
40-50
50-80
80-200
>=200
Query
1
250
500
750
1000
1250
B
GANTC
Protected from DNase
dyad
conserved

Several other authors have suggested that the origin of replication resides within the AT-rich region of the *repC* gene, which happens to be a conserved feature among all *repABC* replicons (Fig. 4.11 and Bartosik *et al.*, 1998; Petersen *et al.*, 2009; Wagner-Dobler *et al.*, 2010). In light of the work done with plasmid p42d, the fact that origins are located inside AT-rich regions and that origins frequently reside within initiator genes, we believe that the RepC binding site described here overlaps the origin of replication.

Our experiments do not reveal the absolute number or stoichiometry of DNA and protein subunits found in each complex. However, binding shows a high degree of cooperativity, with a Hill coefficient of approximately 2 (Fig. 4.4). The simplest interpretation is that RepC multimerizes upon DNA binding. The Hill coefficient, the rather short length of the DNase I protected sequence, and the dyad symmetry of the binding site together provide suggestive evidence that RepC could exist as a monomer in solution and as a dimer when DNA bound. The structural similarity to the MarR family of transcriptional factors also suggests that RepC might exist as a dimer when DNA bound (Fig. 5, Hong *et al.*, 2005). It is also possible that the complexes we detect consist of more than one DNA fragment. Several other plasmids limit replication by the formation of "handcuffed" plasmids, in which the replication initiators bind to the *oriV* of two sister plasmids, preventing further replication until the partitioning system can pull them apart, breaking the handcuffs (Chattoraj, 2000).

Several replication initiator proteins generally bind to the origin, one subunit per iteron repeat. These initiators usually recruit DnaA, and binding of the initiator and DnaA to *oriV* is often cooperative. This sometimes causes DNA looping, which generally induces melting of the DNA strands at the AT-rich region, similar to what happens at *oriC* in *E. coli*. The Rep-DNA complex, in association with DnaA, recruits the replicative helicase (DnaB in *E. coli*) to the bubble created in the AT-rich region

(del Solar *et al.*, 1998). The fact that no iterative sequences are found anywhere in *repC*, and in light of the rather short region bound by RepC, a DNA loop seems improbable. It nonetheless seems plausible that RepC could recruit the replicative helicase with the help of DnaA. Even though a consensus DnaA binding site cannot be found anywhere near the RepC binding site, it has been shown that DnaA can bind to more relaxed consensus sequences (Messer, 2002), therefore this possibility remains to be tested. We have also noticed that a region of DNA directly adjacent to the RepC binding site is extremely conserved among *repC* genes (Fig. 4.11). Given that RepC does not bind to this sequence, at least in *vitro*, we hypothesize that it could serve as a binding site for another, yet to be identified protein.

Plasmids must use host cellular components to replicate. It seems quite puzzling that many of the *repABC* replicons, including the Ti plasmid, carry a copy of the *dnaE* gene, which codes for a homolog of the α subunit of the DNA polymerase III (Slater *et al.*, 2009; Wood *et al.*, 2001). The maintenance of mini-*repC* replicons does not require this gene, as a region containing only the *repC* suffices for plasmid replication (present work). However *dnaE* might have a potential secondary role in plasmid replication, possibly aiding in the replication of the multipartite genomes of many *Agrobacterium* species. This role needs to be addressed in future studies.

We are particularly interested in the GANTC sites at or near *oriV* as well as in the *repE* promoter, and right upstream of the *repA* start codon, especially in light of the studies of GANTC methylation in *C. crescentus* and its role in the cell cycle. In that organism, GANTC sites are found at the chromosomal replication origin and in several promoters that are up- or down-regulated by methylation. The CcrM methylase is synthesized at one point in the cell cycle and only then its hemimethylated DNA is converted to a fully methylated form (Collier *et al.*, 2007). Many copies of a similar sequence, GATC, are found at the replication origin of *E.*

coli. Methylation of these sites by Dam is slowed by the binding of SeqA, which binds hemimethylated DNA but not fully methylated DNA. Binding of SeqA protects the origin from the initiator DnaA, delaying the next round of replication (Katayama *et al.*, 2010).

We were initially disappointed by the finding that methylation did not affect the binding affinity of RepC. In hindsight, this is the result we should have expected. By analogy with *E. coli* it seems plausible that a protein analogous to SeqA could bind to hemimethylated DNA at the *oriV*, and that binding would block RepC binding. SeqA homologs are found throughout the gamma-proteobacteria, but are not found outside this group. However, an analogous protein could bind at each of the GANTC sites at or near the binding site, forming a complex similar to that observed between SeqA and the *E. coli* origin (Waldminghaus and Skarstad, 2009).

The structure of part of RepC can be modeled using the structure of OhrR-DNA complexes (Fig. 4.6) (Hong *et al.*, 2005). OhrR is a dimer that binds to a dyad symmetrical DNA sequence (TACAATTAATTGTA) (Fuangthong *et al.*, 2001). Each OhrR subunit has a recognition helix that binds to one DNA half-site. The two half-sites have a size and spacing similar to the AATGAGTCCCATT site of RepC. We note that the RepC binding site has the sequence GAGTC, a methylation site. Using the OhrR-DNA structure as a model for the RepC-DNA complex, we predict that the methylated A residues of each strand project outwards on the face opposite bound by OhrR. Therefore, we should have predicted that methylation would not alter binding affinity for RepC. The idea that hemimethylation of this site might cause binding of a replication inhibitor, analogous to SeqA, predicts that alteration of this or other methylation sites would cause an increase in plasmid copy number. This hypothesis is currently being tested.

We hoped to detect differences in binding affinities by adding ATP to the binding reactions as it was shown for RepA protein (Pappas and Winans, 2003b), however no difference was observed. Like other replication initiator proteins from iteron plasmids, RepC does not have an ATP binding domain (Chattoraj, 2000). As pointed out by this author, these proteins might help to organize the origin region in order to promote local DNA melting by some other factor(s).

4.5. Experimental Procedures

Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table 4.4. *E. coli* strains were cultured in Luria-Bertani broth (LB) or solid medium at 37°C (Miller, 1972). *A. tumefaciens* strains were cultured in liquid or solid AT minimal medium at 28°C (Cangelosi et al., 1991) or in LB. Antibiotics were added at the following concentrations: 100 μg ml^{-1} spectinomycin; 100 μg ml^{-1} kanamycin, and 100 μg ml^{-1} ampicillin for *E. coli* strains, and 100 μg ml^{-1} of spectinomycin and kanamycin, and 50 μg ml^{-1} of carbenicilin for *Agrobacterium* strains. IPTG was added at the indicated concentrations.

DNA manipulations

Recombinant DNA techniques were performed using established procedures (Sambrook and Russel, 2001). Plasmid DNA was isolated from *E. coli* cells using QIAprep spin miniprep kits (Qiagen). DNA fragments generated by PCR or restriction digestion were gel purified using GeneJET purification columns (Fermentas). Restriction endonucleases were obtained from New England Biolabs and used according to methods described by the manufacturers. Plasmids were introduced into *E. coli* by transformation and into *A. tumefaciens* strains by electroporation

(Cangelosi *et al.*, 1991). Oligonucleotides were purchased from IDT (Integrated DNA technologies) and are described in Table 4.5. Site directed mutagenesis was performed as previously described (Cho *et al.*, 2009). Quantification of DNA was done by using a NanoDrop (Thermo Scientific).

Plasmid isolation from *A. tumefaciens*

A. tumefaciens cells were grown on 20 ml of LB medium supplemented with appropriate antibiotics and were harvested at an OD600 of 0.6. Cells were washed with 4 ml of a buffer composed of 50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 20 mM EDTA, 0.05% Na- Sarkosyl, centrifuged and the pellet resuspended in 400 ul of solution I supplemented with 5 mg/ml of lysozyme, followed by incubation at room temperature for 5 min (Chai and Winans, 2005a). 800 ul of Solution II and 600 ul of Solution III (Sambrook and Russel, 2001) were added sequentially, mixed gently, and incubated on ice for 5 min during each step. Mixtures were cleared by centrifugation at 14,000 g for 10 min and supernatants were precipitated by using 2 volumes of 95% ethanol, followed by incubation at -20°C for 30 min. Samples were centrifuged for 15 min at 14000 RPM at room temperature. The pellets were washed with 300 ul of ethanol 75%, dried at room temperature for 30 min, and resuspended in 250 ul of distilled water. DNA was further purified using GeneJET purification columns (Fermentas) following the provided protocol. DNA was finally eluted in 40 μl of water pH7.7. Plasmid DNA was linearized with appropriate restriction enzymes as indicated and size-fractionated on 1% agarose gel for 30 min to 2 h at 80V/cm.

Table 4.4. Strains and plasmids used in this study.

Strains	Relevant features	References
DH5α	*E. coli*, α-complementation	Stratagene
BL21/DE3	*E. coli B* Plac-gene 1 of bacteriophage T7	(Studier *et al.*, 1990)
UIA143	*A. tumefaciens* C58 strain, Ti plasmid-less, *recA*–, EryR	(Farrand *et al.*, 1989)
Plasmids		
pBluescript SK+	Cloning vector, ColE1 ori, Ap^R	Stratagene
pPZP201	Broad-host-range cloning vector, Sp^R	(Hajdukiewicz *et al.*, 1994)
pSRKKm	Broad-host-range cloning vector ,Kam^R, $lacI^q$	(Khan *et al.*, 2008)
pMCSG9	PT7-his6-MBP-TEV, Ap^R	(Donnelly *et al.*, 2006)
pMalC2	Ptac-malE, colE1 ori, Ap^R	New England biolabs
pSW208	Ori-p15A, Cm^R	(Habeeb *et al.*, 1991)
pKP23	*repABC* operon cloned into a suicide vector, Ap^R	(Pappas and Winans, 2003a)
pYC183	*repE-repC* cloned into pBluescript SK+	(Chai and Winans, 2005b)
pYC189	ATG less *repC* cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pYC159	*repC* cloned into SspI site of pMCSG9 by using ligation-independent-clonging	This study
pUP001	pPZP201 *Age*I/*Sca*I 3666 bp fragment. Suicide vector for *Agrobacterium*, Sp^R	This study
pUP003	pUP001 containing a PLac-*repC* fusion. It was constructed by digesting pUP455 with *Age*I/*Sca*I, and self ligating the 4960 bp fragment that contains the P*lac-repC* fusion and the Sp^R cassette	This study
pUP004	pUP003 digested with *Ase*I/*Bam*HI. Promoter-less *repC*	This study
pUP005	Frameshift mutation on *repC* after codon 317 of *lacZ-repC* fusion on pUP003	This study
pUP131	ATG less *repC* (codons 1-200) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP133	ATG less *repC* (codons 1-320) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP138	ATG less *repC* (codons 1-171) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study

Table 4.4. (continued)

pUP139	ATG less *repC* (codons 30-171) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP140	ATG less *repC* (codons 60-171) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP462	ATG less *repC* (codons 3-171) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP465	ATG less *repC* (codons 19-171) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP468	ATG less *repC* (codons 9-171) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP470	ATG less *repC* (codons 26-171) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP467	ATG less *repC* (codons 1-161) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP472	ATG less *repC* (codons 1-158) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2.	This study
pUP466	ATG less *repC* (codons 1-152) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP132	ATG less *repC* (codons 200-440) cloned in frame with *malE* into the *Bam*HI/*Hind*III sites of pMalC2	This study
pUP450	*repC* cloned into pPZP201. repC original ribosomal binding site	This study
pUP455	*repC* cloned into pPZP201, translational fusion between *lacZ*α-*repC*	This study
pUP500	*repC* cloned into NdeI/*Hind*III sites of pSRKKm	This study
pUP501	*ccrM* cloned into NdeI/*Hind*III sites of pSRKKm	This study
pUP502	pYC183 digested with NheI and filled in with T4 polymerase, religated to generate a RepC truncated (un-functional) protein after codon 87	This study
pUP509	pYC183 with a frameshift mutation on codon 303 of *repC* generated by site directed mutagenesis	This study

Table 4.5. Oligonucleotides used in this study.

Fragment number	Oligonucleotide name	DNA sequence
For the *repC* fragments used in the gel shift assays		
Fragment1	Set1fwd*Sac*I	GACGAGCTCTGGATCCGATGAAATGAACCG
Fragment1	Set1rev*Kpn*I	CGGGGTACCCACGAGAACTGCAAGATGCCG
Fragment2	Set1fwd*Sac*I	GACGAGCTCTGGATCCGATGAAATGAACCG
Fragment2	5'Rev2	GCTGGTACCTCTGCATTTCCTTGTACC
Fragment3	For-igs2-*Sac*I	GCTGAGCTCGAAAGTATCGTTTCGACG
Fragment3	5'Rev3	TATTCTGGAACACATGCC
Fragment4	Set2fwd*Sac*I	GACGAGCTCGCAAAGGCCGTGGCG
Fragment4	Set1rev*Kpn*I	CGGGGTACCCACGAGAACTGCAAGATGCCG
Fragment5	Set2fwd*Sac*I	GACGAGCTCGCAAAGGCCGTGGCG
Fragment5	Set2rev*Kpn*I	CGGGGTACCAACGCCCTCTTCGAT
Fragment6	Set3fwd*Sac*I	GACGAGCTCGAGTGCGGCCTGGTCATC
Fragment6	Set3rev*Kpn*I	CGGGGTACCTGGTTTATTTCCAAAGCCGTA
Fragment7	Set4fwd*Sac*I	GACGAGCTCCCCGGCAACTGGTGC
Fragment7	Set4rev*Kpn*I	CGGGGTACCCGCCGCAAGCGTGATT
Fragment8	Set5fwd*Sac*I	GACGAGCTCGAAGCGGGCGGCACG
Fragment8	Set5rev*Kpn*I	CGGGGTACCTATTCATTTCGGCCTCGACTT
Fragment8B*	Set6fwd*Bam*HI	CGCGGATCCATCTATCAGAAGGCCGAC
Fragment8B*	Rev PstI repC Set6	GCTCTGCAGACAGGGCGCGTCCCATTCG
Primers used to narrow down the binding site		
Fragment9	Set3fwd*Sac*I with Set2rev*Kpn*I	
Fragment10	Set4fwd*Sac*I with Set3rev*Kpn*I	
Fragment11	Set5fwd*Sac*I with Set4rev*Kpn*I	
For fragments 12 to 15 the forward primer was Set4fwd*Sac*I		
Fragment12	Set3bot*Kpn*Ib	CGGGGTACCTTCAGATTCAGGTATAGAGTC
Fragment13	Set3bot*Kpn*Ic	CGGGGTACCTGGTTTTGAATTCTGTATGTG
Fragment14	Set3bot*Kpn*Id	CGGGGTACCACGACCAAAATGGGACTCATT
Fragment15	Set3bot*Kpn*Ie	CGGGGTACCTGGCGTCCTGAATCTGTG

Table 4.5. (continued)

For fragments 16 to 20 the reverse primer was Set3rev*Kpn*I		
Fragment16	Set4fwd*Sac*Ib	GACGAGCTCGTGCAGCAGGTGTATCAAGCG
Fragment17	Set4fwd*Sac*Id	GACGAGCTCCACGCTCCGCACCAAGACAGC
Fragment18	Set4ForF2	AAATCCGTGACGTCTTGGAAT
Fragment19	Set4ForH	AGGACGCCAATGAGTCCCATT
Fragment20	Set4ForI	TTGGTCGTCACATACAGAATT
Primers for the methylation studies		
	Met1-for	TTCAGGACGCCAATGXGTCCCATTTTGGTCG TCACATACAGA
	Met1-rev	TCTGTATGTGACGACCAAAATGGGXCTCATT GGCGTCCTGAA
Primer for in vitro footprinting. Used together with Set4fwd*Sac*I		
	FAM-Set3 Bottom primer	/56-FAM/TGGTTTATTTCCAAAGCCG
Primers with mutations in the origin		
	Ori1M-For	TACAAGATGCTAACGAAAGTCACTTCGGC AGACATATCCAGA
	Ori1M-Rev	TCTGGATATGTCTGCCGAAGTGACTTTCGT TAGCATCTTGTA
	F1repC	GGGAACAAAAGCTGGAGC
	R1-repC	AGGGCGAATTGGGTACCG
	Mut-Ori-vivoF	ACAAGATGCTAACGAAAGTCACTTCGGCA GACATATCCAGAATTCAAAACC
	Mut-Ori-vivoR	GATATGTCTGCCGAAGTGACTTTCGTTAGC ATCTTGTATCTGTGTTTTTGC
For pUP500 construction		
	For-repC-*Nde*I	CGTCATATGCAGACGCATTTATC
	Rev-repC440-*Hind*III	GCTAAGCTTTCATTTCGGCCTCGAC
For pUP450 and 455 construction		
pUP450	ForNTD-*Bam*HI-pPZP201	GTCGGATCCAGTATCGTTTCGACGAGC
	Rev-repC440-*Hind*III	GCTAAGCTTTCATTTCGGCCTCGAC
pUP455	ForNTD2-*Bam*HI-pPZP201	GTCGGATCCACAGACGCATTTATCAAC

Table 4.5. (continued)

	Rev-repC440-*Hind*III	GCTAAGCTTTCATTTCGGCCTCGAC
For construction of RepC protein truncations on pMalC2		
For pUP133 to 466 the for primer was for-repC *Bam*HI		
	For-repC *Bam*HI	GAATTCGGATCCCAGACG
pUP133	Rev-repC320*Hind*III	GCTAAGCTTTCAATCGCGCCAATGGCGGAT
pUP131	Rev-repC200*Hind*III	GCTAAGCTTTCACACCTGCTGCACTCC
pUP138	Rev-repC170*Hind*III	GCTAAGCTTTCACAGCCGCTCCTTCGCGAC
pUP467	Rev-repC160*Hind*III	GCTAAGCTTTCATTTCTTCTCAGCCTGGATC
pUP472	Rev-repC158*Hind*III	GCTAAGCTTTCAAGCCTGGATCGCTTCGGC
pUP466	Rev-repC150*Hind*III	GCTAAGCTTTCACATATCTCTGAATTCTTTGG
For pUP462 to 140 the reverse primer was Rev-repC170-*Hind*III		
pUP462	For-repCD3 *Bam*HI	GCTGGATCCTTATCAACGACGCCCTTTG
pUP468	For-repCD10 *Bam*HI	GCTGGATCCGGGCGGCGGCCGATGACTC
pUP465	For-repCD20 *Bam*HI	GCTGGATCCTCAAGTCAGATGTCAGC
pUP470	For-repCD26*Bam*HI	GCTGGATCCAAGGCCGTGGCGCCTGAC
pUP139	For-repC-D30-*Bam*HI	GCTGGATCCGCGCCTGACGCTACCGCA
pUP140	ForRepC D60 *Bam*HI	GCTGGATCCGCGATCCTCAATGCCTTG
For pUP132 the reverse primer was Rev-repC440-*Hind*III		
pUP132	For-repC200 *Bam*HI	TTCGGATCCGTGTATCAAGCGATTATC
	Rev-repC440-*Hind*III	GCTAAGCTTTCATTTCGGCCTCGAC
Primers used to construct pYC189		
	RepF12	GTACGGATCCCAGACGCATTTATCAACGAC
	Rep R3	GCTAAGCTTAACCGGATATTCATTTCGGC
Primers used to construct pYC159		
	RepC F2LIC	TATTTCCAATCCAATGCAATGCAGACGCA TTTATCA
	RepC R4LIC	TTATCCACTTCCAATGTCATTTCGGCCTC GACTT
Primers used to construct frameshift mutation at 3’ end of *repC* for plasmids pUP005 and pUP509		
	Fs1-for	CTGAATGCCTaGCCCGAGCGTG

Table 4.5. (continued)

Fs1-rev	CACGCTCGGGCtAGGCATTCAG
For 003	TGTGGAATTGTGAGCGG
Rev 003	GATTAAGTTGGGTAACGC

X is a methylated adenine (2-Amino-2'-dA) constructed by IBA nucleic Acids Synthesis (Germany).
* Fragment8B overlaps with Fragment 8, and extends downstream of the *repC* stop codon into the vector pBSK+. This fragment was not shifted by ReC.

Protein purification

BL21(DE3) strains containing plasmids with various *malE-repC* fusions were cultured at 37°C in 300 ml of LB broth containing 200 ug of ampicillin per ml. When cultures reached an OD600 nm of 0.4, 500 uM of IPTG were added to the cultures together with 200 ug per ml of ampicillin. Incubation was continued at 37°C for 2.5 h. Cells were collected, resuspended in 5 ml of TEDG buffer (Tris-HCl 50 mM pH 8.0, EDTA 0.5 mM, DTT 1 mM, glycerol 10%) supplemented with 200 mM NaCl, and disrupted twice using a French pressure minicell (15,000 lb/in^2). Total cell lysates were separated into soluble and pellet fractions by centrifugation at 14,000 g for 30 min at 4°C. Clear supernatants were applied to an amylose affinity chromatography column (New England Biolabs). Proteins were step eluted with buffer TEDG with 200 mM NaCl containing 15 mM maltose. Samples were run on 12% SDS-PAGE gels and fractions containing the protein were dialyzed overnight at 4°C against TEDG buffer containing 50% glycerol and 200 mM NaCl. Protein preparations were aliquoted and stored at -80°C.

In order to separate RepC from MBP, overexpression was performed from strain BL21/DE3(pYC159). The His-MBP-RepC fusion was obtained by amylose affinity chromatography. Pooled fractions containing the fusion protein were dialyzed against buffer phosphate 50 mM pH 7.4 + glycerol 5% + NaCl 100 mM. Purified

fusion protein was cut with His-TEV protease on an overnight digestion at 4°C. The concentration of TEV protease that gave best results was 3 OD_{280} of TEV protease per 100 OD_{280} of His-MBP-RepC. The digestion reaction was applied to a Nickel column (Ni Sepharose 6 Fast Flow, Amersham Biosciences) with recirculation of the sample for 15 min at RT, and RepC was collected in the flow through since His-MBP and His-TEV proteins are trapped in the column.

Electrophoretic mobility shift assay (EMSA)

DNA fragments were PCR amplified using oligos described in Table 4.5. End-labelling reaction was performed with [γ-32P]-dATP (Pelkin Elmer) using T4 polynucleotide kinase (New England Biolabs). Negative control DNA fragment was amplified by using primers previously described (Pinto and Winans, 2009). Binding reactions containing 10^{-12} M DNA and proteins in various amounts as indicated were conducted as previously described (Pinto and Winans, 2009). Gels were analyzed by using a Storm PhosphorImager B840 (Molecular Dynamics).

DNase I protection assay

A 250 bp *repC* fragment was PCR amplified with primers Set4fwd*Sac*I and FAM-Set3 Bottom primers (Table 4.5) using pYC183 as a template. The experiment was done as described by (Zianni *et al.*, 2006) in three repetitions with some modifications. 2 μl of DNA (~200 ng) was diluted in 20 μl of gel shift buffer (2X) (Pinto and Winans, 2009); then 5 μl of MBP-RepC (1mg/ml) (or BSA for control) and 13 μl of water were added to the reaction; binding was allowed to proceed for 30 min on ice; then 0.8 μl of $MgCl_2$/$CaCl_2$ mix (50X) were added; tubes were incubated at RT for 5 min to equilibrate, and 0.1U of DNaseI was added to the reaction with incubation for 30 min at RT; the enzyme was heat inactivated on PCR machine at 95°C for 5 min;

DNA was purified by using Qiagen PCR kit and eluted on 25 μl of water. Fragments were sent to the Cornell University Life Sciences Core Laboratories Center for DNA fragment analysis using the Applied BioSystems 3730xl DNA Analyzer.

Replication *in cis* experiment

In order to test whether RepC works preferentially *in cis* or *in trans*, we performed an experiment by taking advantage of RepC expressed from the *lac* promoter in two different fusions. In the first fusion, *repC* and 50 bp of upstream sequence containing the original *repC* ribosomal binding site (RBS) were PCR amplified and cloned into the *Bam*HI/*Hind*III sites of pPZP201 creating a transcriptional fusion with the Plac promoter, this plasmid was called pUP450. In the second fusion, *repC* was cloned by using the same restriction enzymes, but in frame with the first 14 amino acids of the LacZα peptide, creating a translational fusion with the strong RBS of the *lacZα* gene. This plasmid was called pUP455.

Plasmid pUP450 or pPZP201 (vector control) were electroporated onto UIA143 strain containing plasmid pYC183 or pKP23. Both pYC183 and pKP23 depend on RepC and its origin for replication. Plasmids were extracted as above described and plasmid preparations were subjected to restriction digestion as indicated.

Plasmid pUP450, pUP455 or pPZP201 were electroporated onto UIA143 strain containing pSRKKm (the source for the LacI repressor, Table 4.4) and pKP23. Cultures were grown with appropriate antibiotics and IPTG was added as indicated. Plasmids were extracted as described above and plasmid preparations were subjected to restriction digestion as indicated.

The digestion pattern from these plasmids is as follows:

*Kpn*I and *Hind*III: pPZP201 – 7.1 kb; pUP450 – 7.1 and 1.3 kb; pKP23 – 5.9, 3, 2, and 1 kb; pYC183 – 4.4 kb.

*Kpn*I: pPZP201. – 7.1 kb; pUP455 – 8.5 kb; pSRKKm – 5.8 kb.

*Hind*III: pPZP201. – 7.1 kb; pUP455 – 8.5 kb; pSRKKm – 5.8 kb; pKP23 – 5.9, 4, and 2 kb.

Acknowledgments

The authors gratefully acknowledge Dr. Yunrong Chai (Harvard University) for construction of plasmids pYC159 and pYC189. This study was supported by a grant from the NIGMS (GM042893). U.M.P. acknowledges the partial financial support from the Brazilian government through a fellowship grant from the Coordenação de Aperfeicoamento de Pessoal de Nível Superior (Capes). We are also grateful to our laboratory members for the critical reading of the manuscript.

REFERENCES

Alvarez-Martinez, C.E., and Christie, P.J. (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**: 775-808.

Bartosik, D., Baj, J., and Wlodarczyk, M. (1998) Molecular and functional analysis of pTAV320, a repABC-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. *Microbiology* **144**: 3149-3157.

Bingle, L.E., and Thomas, C.M. (2001) Regulatory circuits for plasmid survival. *Curr Opin Microbiol* **4**: 194-200.

Bruand, C., and Ehrlich, S.D. (1998) Transcription-driven DNA replication of plasmid pAMbeta1 in *Bacillus subtilis*. *Mol Microbiol* **30**: 135-145.

Cangelosi, G.A., Best, E.A., Martinetti, G., and Nester, E.W. (1991) Genetic analysis of *Agrobacterium*. *Methods Enzymol* **204**: 384-397.

Castillo-Ramirez, S., Vazquez-Castellanos, J.F., Gonzalez, V., and Cevallos, M.A. (2009) Horizontal gene transfer and diverse functional constrains within a common replication-partitioning system in Alphaproteobacteria: the repABC operon. *BMC Genomics* **10**: 536.

Cervantes-Rivera, R., Romero-Lopez, C., Berzal-Herranz, A., and Cevallos, M.A. (2010) Analysis of the mechanism of action of the antisense RNA that controls the replication of the *repABC* plasmid p42d. *J Bacteriol* **192**: 3268-3278.

Cevallos, M.A., Cervantes-Rivera, R., and Gutierrez-Rios, R.M. (2008) The *repABC* plasmid family. *Plasmid* **60**: 19-37.

Chai, Y., and Winans, S.C. (2005a) RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between repA and repB for plasmid partitioning and autorepression. *Mol Microbiol* **58**: 1114-1129.

Chai, Y., and Winans, S.C. (2005b) A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. *Mol Microbiol* **56**: 1574-1585.

Chattoraj, D.K. (2000) Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol Microbiol* **37**: 467-476.

Chen, C.Y., and Winans, S.C. (1991) Controlled expression of the transcriptional activator gene virG in *Agrobacterium tumefaciens* by using the *Escherichia coli lac* promoter. *J Bacteriol* **173**: 1139-1144.

Cho, H., and Winans, S.C. (2005) VirA and VirG activate the Ti plasmid repABC operon, elevating plasmid copy number in response to wound-released chemical signals. *Proc Natl Acad Sci U S A* **102**: 14843-14848.

Cho, H., and Winans, S.C. (2007) TraA, TraC and TraD autorepress two divergent quorum-regulated promoters near the transfer origin of the Ti plasmid of *Agrobacterium tumefaciens*. *Mol Microbiol* **63**: 1769-1782.

Cho, H., Pinto, U.M., and Winans, S.C. (2009) Transsexuality in the rhizosphere: quorum sensing reversibly converts *Agrobacterium tumefaciens* from phenotypically female to male. *J Bacteriol* **191**: 3375-3383.

Collier, J., McAdams, H.H., and Shapiro, L. (2007) A DNA methylation ratchet governs progression through a bacterial cell cycle. *Proc Natl Acad Sci U S A* **104**: 17111-17116.

Crossman, L.C. (2005) Plasmid replicons of Rhizobium. *Biochem Soc Trans* **33**: 157-158.

Das, N., and Chattoraj, D.K. (2004) Origin pairing ('handcuffing') and unpairing in the control of P1 plasmid replication. *Mol Microbiol* **54**: 836-849.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M., and Diaz-Orejas, R. (1998) Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**: 434-464.

Derbyshire, K.M., and Grindley, N.D. (1996) Cis preference of the IS903 transposase is mediated by a combination of transposase instability and inefficient translation. *Mol Microbiol* **21**: 1261-1272.

Dessaux, Y., Petit, A., Farrand, S.K., and Murphy, P.J. (1998) Opines and opine-like molecules involved in plant-*Rhizobiaceae* interactions. In *The Rhizobiaceae*. Spaink, H.P., Kondorosk, A. and Hooykaas, P.J. (eds). Dordrecht: Kluwer Academic Publishers, pp. 173-197.

Donnelly, M.I., Zhou, M., Millard, C.S., Clancy, S., Stols, L., Eschenfeldt, W.H., Collart, F.R., and Joachimiak, A. (2006) An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expr Purif* **47**: 446-454.

Farrand, S.K., O'Morchoe, S.P., and McCutchan, J. (1989) Construction of an *Agrobacterium tumefaciens* C58 recA mutant. *J Bacteriol* **171**: 5314-5321.

Francia, M.V., Fujimoto, S., Tille, P., Weaver, K.E., and Clewell, D.B. (2004) Replication of *Enterococcus faecalis* pheromone-responding plasmid pAD1: location of the minimal replicon and oriV site and RepA involvement in initiation of replication. *J Bacteriol* **186**: 5003-5016.

Fuangthong, M., Atichartpongkul, S., Mongkolsuk, S., and Helmann, J.D. (2001) OhrR is a repressor of ohrA, a key organic hydroperoxide resistance determinant in *Bacillus subtilis*. *J Bacteriol* **183**: 4134-4141.

Funnell, B.E. (2005) Partition-mediated plasmid pairing. *Plasmid* **53**: 119-125.

Gering, M., Gotz, F., and Bruckner, R. (1996) Sequence and analysis of the replication region of the *Staphylococcus xylosus* plasmid pSX267. *Gene* **182**: 117-122.

Habeeb, L.F., Wang, L., and Winans, S.C. (1991) Transcription of the octopine catabolism operon of the *Agrobacterium* tumor-inducing plasmid pTiA6 is activated by a LysR-type regulatory protein. *Mol Plant Microbe Interact* **4**: 379-385.

Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol Biol* **25**: 989-994.

Hong, M., Fuangthong, M., Helmann, J.D., and Brennan, R.G. (2005) Structure of an OhrR-ohrA operator complex reveals the DNA binding mechanism of the MarR family. *Mol Cell* **20**: 131-141.

Izquierdo, J., Venkova-Canova, T., Ramirez-Romero, M.A., Tellez-Sosa, J., Hernandez-Lucas, I., Sanjuan, J., and Cevallos, M.A. (2005) An antisense RNA plays a central role in the replication control of a *repC* plasmid. *Plasmid* **54**: 259-277.

Johnson, T.J., and Nolan, L.K. (2009) Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol Mol Biol Rev* **73**: 750-774.

Kahng, L.S., and Shapiro, L. (2003) Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. *J Bacteriol* **185**: 3384-3391.

Katayama, T., Ozaki, S., Keyamura, K., and Fujimitsu, K. (2010) Regulation of the replication cycle: conserved and diverse regulatory systems for DnaA and *oriC*. *Nat Rev Microbiol* **8**: 163-170.

Kelley, L.A., and Sternberg, M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**: 363-371.

Khan, S.R., Gaines, J., Roop, R.M., 2nd, and Farrand, S.K. (2008) Broad-host-range expression vectors with tightly regulated promoters and their use to examine

the influence of TraR and TraM expression on Ti plasmid quorum sensing. *Appl Environ Microbiol* **74**: 5053-5062.

Kwong, S.M., Skurray, R.A., and Firth, N. (2004) *Staphylococcus aureus* multiresistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense RNA regulation. *Mol Microbiol* **51**: 497-509.

Li, P.L., and Farrand, S.K. (2000) The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. *J Bacteriol* **182**: 179-188.

MacLellan, S.R., Smallbone, L.A., Sibley, C.D., and Finan, T.M. (2005) The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. *Mol Microbiol* **55**: 611-623.

McAdams, H.H., and Shapiro, L. (2003) A bacterial cell-cycle regulatory network operating in time and space. *Science* **301**: 1874-1877.

McFall, E. (1986) cis-acting proteins. *J Bacteriol* **167**: 429-432.

Messer, W. (2002) The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev* **26**: 355-374.

Nikaido, H. (2009) Multidrug resistance in bacteria. *Annu Rev Biochem* **78**: 119-146.

Ozaki, S., and Katayama, T. (2009) DnaA structure, function, and dynamics in the initiation at the chromosomal origin. *Plasmid* **62**: 71-82.

Pappas, K.M., and Winans, S.C. (2003a) A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. *Mol Microbiol* **48**: 1059-1073.

Pappas, K.M., and Winans, S.C. (2003b) The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. *Mol Microbiol* **49**: 441-455.

Pappas, K.M. (2008) Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. *Plasmid* **60**: 89-107.

Paulsson, J., and Chattoraj, D.K. (2006) Origin inactivation in bacterial DNA replication control. *Mol Microbiol* **61**: 9-15.

Petersen, J., Brinkmann, H., and Pradella, S. (2009) Diversity and evolution of *repABC* type plasmids in *Rhodobacterales*. *Environ Microbiol* **11**: 2627-2638.

Phale, P.S., Basu, A., Majhi, P.D., Deveryshetty, J., Vamsee-Krishna, C., and Shrivastava, R. (2007) Metabolic diversity in bacterial degradation of aromatic compounds. *Omics* **11**: 252-279.

Pinto, U.M., and Winans, S.C. (2009) Dimerization of the quorum-sensing transcription factor TraR enhances resistance to cytoplasmic proteolysis. *Mol Microbiol* **73**: 32-42.

Ravin, N.V., Kuprianov, V.V., Gilcrease, E.B., and Casjens, S.R. (2003) Bidirectional replication from an internal ori site of the linear N15 plasmid prophage. *Nucleic Acids Res* **31**: 6552-6560.

Sambrook, J., and Russel, D.W. (2001) *Molecular cloning - a Laboratory manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Scherer, G. (1978) Nucleotide sequence of the O gene and of the origin of replication in bacteriophage lambda DNA. *Nucleic Acids Res* **5**: 3141-3156.

Slater, S.C., Goldman, B.S., Goodner, B., Setubal, J.C., Farrand, S.K., Nester, E.W., Burr, T.J., Banta, L., Dickerman, A.W., Paulsen, I., Otten, L., Suen, G., Welch, R., Almeida, N.F., Arnold, F., Burton, O.T., Du, Z., Ewing, A., Godsy, E., Heisel, S., Houmiel, K.L., Jhaveri, J., Lu, J., Miller, N.M., Norton, S., Chen, Q., Phoolcharoen, W., Ohlin, V., Ondrusek, D., Pride, N., Stricklin, S.L., Sun, J., Wheeler, C., Wilson, L., Zhu, H., and Wood, D.W. (2009) Genome

sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol* **191**: 2501-2511.

Smith, E.F., and Townsend, C.O. (1907) A plant-tumor of bacterial origin. *Science* **25**: 671-673.

Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**: 60-89.

Tanaka, T., Ishida, H., and Maehara, T. (2005) Characterization of the replication region of plasmid pLS32 from the Natto strain of *Bacillus subtilis*. *J Bacteriol* **187**: 4315-4326.

Thomas, C.M. (2000) Paradigms of plasmid organization. *Mol Microbiol* **37**: 485-491.

Venkova-Canova, T., Soberon, N.E., Ramirez-Romero, M.A., and Cevallos, M.A. (2004) Two discrete elements are required for the replication of a *repABC* plasmid: an antisense RNA and a stem-loop structure. *Mol Microbiol* **54**: 1431-1444.

Wagner-Dobler, I., Ballhausen, B., Berger, M., Brinkhoff, T., Buchholz, I., Bunk, B., Cypionka, H., Daniel, R., Drepper, T., Gerdts, G., Hahnke, S., Han, C., Jahn, D., Kalhoefer, D., Kiss, H., Klenk, H.P., Kyrpides, N., Liebl, W., Liesegang, H., Meincke, L., Pati, A., Petersen, J., Piekarski, T., Pommerenke, C., Pradella, S., Pukall, R., Rabus, R., Stackebrandt, E., Thole, S., Thompson, L., Tielen, P., Tomasch, J., von Jan, M., Wanphrut, N., Wichels, A., Zech, H., and Simon, M. (2010) The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. *Isme J* **4**: 61-77.

Waldminghaus, T., and Skarstad, K. (2009) The *Escherichia coli* SeqA protein. *Plasmid* **61**: 141-150.

Wood, D.W., Setubal, J.C., Kaul, R., Monks, D.E., Kitajima, J.P., Okura, V.K., Zhou, Y., Chen, L., Wood, G.E., Almeida, N.F., Jr., Woo, L., Chen, Y., Paulsen, I.T., Eisen, J.A., Karp, P.D., Bovee, D., Sr., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M.J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Romero, P., Gordon, D., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z.Y., Dolan, M., Chumley, F., Tingey, S.V., Tomb, J.F., Gordon, M.P., Olson, M.V., and Nester, E.W. (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317-2323.

Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J., Farrand, S.K., and Winans, S.C. (2000) The bases of crown gall tumorigenesis. *J Bacteriol* **182**: 3885-3895.

Zianni, M., Tessanne, K., Merighi, M., Laguna, R., and Tabita, F.R. (2006) Identification of the DNA bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. *J Biomol Tech* **17**: 103-113.

Zzaman, S., and Bastia, D. (2005) Oligomeric initiator protein-mediated DNA looping negatively regulates plasmid replication in vitro by preventing origin melting. *Mol Cell* **20**: 833-843.